

25 June 1981

Congress and cancer research

Has the United States Congress, having declared "war on cancer" ten years ago, now decided to declare war on the National Cancer Institute? This speculation has run through the biomedical research community since the extraordinary occasion three weeks ago when the director of the institute, Dr Vincent T. DeVita, was hauled over the coals by the Senate Committee on Labor and Human Resources (see *Nature*, 11 June, p.444), one of the several committees in Congress which have in the past decade cheered on the institute's growth. The committee was the platform from which, in the early 1970s, Senator Edward M. Kennedy sought to outbid the then Administration in generosity towards cancer research. Now, Senator Kennedy finds himself in the unfamiliar role of senior minority member of the committee, subject to the tutelage of Senator Orrin G. Hatch, apparently bent on winning his spurs in the Senate by being a conspicuous and good Republican — eliminating "waste", diminishing the "role of government" and so on. But Dr DeVita was not merely the victim of a change of political style.

So much seems implicitly to have been acknowledged by Dr DeVita's colleagues at the National Cancer Institute, who put out a statement last week praising his qualities both as a scientist and an administrator. Everybody, at the institute and elsewhere, is agreed that Dr DeVita's appointment as director on 1 January brought a breath of fresh air to an institution in danger of drifting complacently with the tide of congressional largesse. He commands the professional respect of his colleagues, but he is also energetic and tough. It would be disastrous for one of the most important of all federally supported scientific institutions if the Senate committee, or some other zealous custodian of the public purse, were at this stage needlessly to interfere with Dr DeVita's chance of making a success of the National Cancer Institute. The particular issues with which he was confronted three weeks ago — the discovery of irregularities in the conduct of research contracts let by the institute — are acknowledged to have their origins in relatively ancient history. To the extent that the National Cancer Institute is at all responsible, the blame should rest with Dr DeVita's predecessors. His own fault — which some might think a virtue — at the hearing three weeks ago was that of political innocence. He seemed not fully to have appreciated that the first duty of any public servant before a congressional committee is to affirm that his every waking hour is spent on saving every possible nickel of taxpayers' money. Only when that has been done is it permissible for him to think of prosecuting the task for which his budget has been set aside, "curing" cancer or whatever it may be.

So far, there is no knowing what the new Senate Committee on Labor and Human Resources has at the back of its collective mind. The complaints against the institute brought up by Senator Hatch were hardly new. Some of them had been rehearsed before Representative Albert Gore's subcommittee in the House of Representatives some weeks earlier. But, in any case, why should the National Cancer Institute itself be blamed if one of the several investigators working on a collaborative clinical research contract has been accused of keeping inadequate or even falsified medical records? (The investigation of these allegations by the National Institutes of Health will not be complete until after the summer.) Dr DeVita was embarrassed three weeks ago because he had not blackballed a further grant to the same researcher, on the principle of "innocent until proven guilty". But who, in Congress of all places, would expect a good administrator to behave differently? So why did Senator Hatch hold his inquiry at all, and

then assiduously leak the details of his complaints in advance to the daily press? Can there be so much smoke and no fire?

Time will tell, probably quite quickly. The unexpected resignation of Dr Donald Fredrickson as director of the National Institutes of Health (of which the National Cancer Institute is one) will be a chance for finding out (see page 603). Although it will be necessary to read the tealeaves before knowing what to make of the White House's nomination of a successor to Fredrickson, Senator Hatch's committee will no doubt seize the opportunity for saying where it stands on continued public support for biomedical research and for cancer research in particular. The greatest danger, in the period immediately ahead, is that the naivety that inspired the war on cancer ten years ago will be found still to persist. Already, some members of the Hatch committee have been asking aloud why, after a decade's generous spending, cancer has not yet been "cured". More generally, the persistence of rheumatoid arthritis, cardiovascular disease or even death itself has lent weight to the popular suspicion that good health is not one of the products of research. Congress, more than usually eager to do the work of the Office of Management and Budget, could easily become jaundiced.

Of necessity, neither obduracy nor round robins in support of Dr DeVita will meet the need. More to the point would be a determined attempt to clear up the misunderstanding between Congress and the research community on the function of cancer research. The notion that there is a war to be won is foolish — and springs largely from the politicians. While the National Cancer Institute has direct responsibility for several important programmes directly related to the improvement of various cancer treatments, the most promising part of its work remains its long-term research, only some of which is ever likely to help with the treatment of specific types of cancer. Congress must somehow be helped to appreciate this simple truth. The task would be easier if it were also more widely appreciated in Congress that the basic research programmes of several of the constituent institutes of the National Institutes of Health are essentially part of a cooperative effort, aimed at the understanding of cell biology and physiology that will, on past experience, make possible as yet unknown treatments for possibly unrecognized illnesses. Administratively, that argues for closer relationships between the several institutes and for flexibility in the disposal of their separate budgets. In the past few years, unfortunately, Congress has been pushing in the opposite direction, seeking more direct control over individual institutes. One consequence is that people like Dr DeVita, appointed as scientists, find themselves dealt with in Congress as if they were nothing but accounting officers. The time has come to reverse this tendency.

No Irish science politics

The assumption that science and politics do not mix is only occasionally falsified. The dangers in too rigorous an attachment to the principle is that it may make a nonsense of the practice. This is what the Science Research Council (now even more augustly known as the Science and Engineering Research Council) has just done. For the council has turned down an innocuous proposal that there should in future be an exchange of graduate students between the United Kingdom and the Irish Republic on the modest scale of a dozen or so a year. By all accounts, the council wishes this modification of pure chauvinism to be subsumed in an



even more imaginative scheme — a grand European scheme for the freedom of movement of graduate students within Europe as a whole. The snag, of course, is that the grander scheme will take ages to arrange. Once more, the best is the enemy of the good.

It will not have escaped the attention of the world at large, scientists included, that the United Kingdom government has a serious problem in Northern Ireland, otherwise called Ulster. Some months ago, it was agreed that there should be informal talks on the development of closer relations between the Republic and the United Kingdom between the prime ministers of the two independent states. The talks were known as the "Thatcher-Haughey" talks until the prime minister of Ireland allowed his political future to be put in doubt at last week's general election. By all accounts, the agenda for these talks has been severely practical — how to organize the more economic exchange of electricity and so on. The long-term objective has been to build common institutions between the two principal governments of such strength that the people who live in Ulster will appreciate, whatever their immediate inclinations, that the status of Ulster is not the only problem in the world. Somewhere along the line, it now seems plain, somebody had the wit to recognize that a fuller integration of higher education between the United Kingdom and the Republic would serve as a demonstration to Ulstermen of all persuasions (but for practical purposes there are only two) that the interests of young people and of scholarship can transcend sheer sectarianism. The proposal was that up to a dozen places for graduate students in one country might be made available (subject to a proper scrutiny of academic standards and so on) to would-be graduate students from the other. It has been turned down by the Science and Engineering Research Council.

On the face of things, the explanation is reasonable enough. Why make an exception of the Republic of Ireland when the United Kingdom is part of a much larger community? So instead of an understanding between the two native English-speaking governments in Europe on graduate education, the Science and Engineering Research Council is to embark on an exploration of the possibilities of exchanging graduate students more widely, perhaps within the European Community or more widely still. With luck, there may be a formal proposal for everybody to discuss five years from now, by which time the climate of Anglo-Irish relations will have changed, perhaps even for the worse.

This is why the Science and Engineering Council of the United Kingdom should think again. It could reconsider its refusal of a deal on Anglo-Irish graduate students under the item "Matters arising from the minutes" at its next meeting. It should then feel free to acknowledge that it was needlessly pedantic at its earlier meeting, and that there are some circumstances in which the improvement of the social climate that may be brought about by an apparently administrative chance are unexpectedly great. Briefly, the benefits of some kind of understanding between the United Kingdom and the Republic of Ireland on the education of graduate students could suggest to people on both sides of the border between them that there is an even stronger case for running the two university systems in concert, and that still other forms of collaboration might yield still greater benefits. The problem of Ulster is, after all, a problem that the people most concerned have decided is too important to be talked about, let alone negotiated. To those who live in the United Kingdom (including Ulster) and in the Republic, the cruel irony is daily denial of well-established links. The Science and Engineering Council has fallen into the trap of supposing that links such as these are irrelevant to its grander purpose. In that, even in its own self-interest, it is mistaken.

Luddites of two kinds

Luddites are not especially welcome when everybody's energy is concentrated on making national and other (regional? global?) economies more productive. Yet Luddites have a valuable place even if only as the grains of sand that enable oysters to grow

pearls. Two interesting but disparate clutches of them have made their appearance recently in the United Kingdom — earlier this week, an organization called Cambridge Econometrics, which is an offshoot of the Department of Applied Economics (at the university of the same name) and which has a well-established reputation for the advocacy of left-wing protectionism, published its latest gloomy economic forecast for the United Kingdom, and the Council for Science and Society, a high-level do-gooder organization last month published its latest document, *New Technology: Society, Employment and Skill*. Both arguments bear on the question whether foreseeable technological developments will make most of us happier than we are. Although different in texture and substance, their conclusions are similar — things will get worse before they can get better; and in the lifetimes of those now alive, the sense of things getting worse will predominate.

Against the odds suggested by its past performance, the Council for Science and Society has the more interesting case. Its document is the product of a committee under Professor Howard Rosenbrock of the University of Manchester Institute of Science and Technology which has obviously had the wit to leave the drafting of its report to one person. The result is a sensitive account of the nineteenth century Industrial Revolution followed by an imaginative guess about the next few decades. The most arresting part of this seemingly rambling train of thought is the brief discussion of the meaning of the word productivity and its suggestion that the concept is not by definition virtuous. If increased productivity means that the same number of people produce more goods collectively, there are more for them to share between them. If, however, increased productivity helps business enterprises to be more competitive, the result may be that they choose to cut each others' throats, not in the process increasing the wellbeing of their workers. Nation-states blessed with increasing productivity have a further degree of freedom — they may choose to spend the extra benefits on research and development or perhaps even on quite different things — armaments for example — but are constrained in the malevolence they are able to cause by social (and political) requirements that, for example, even unemployed workers should be paid nearly a living wage. But in whatever context productivity may be measured, there are potential drawbacks for ordinary people; briefly, they may not share fully in the benefits. This thoughtful document is not a tract but a provoking reminder that more automation — almost certain to come about — does not necessarily mean more happiness.

Cambridge Econometrics, as befits the dismal science that has spawned it, has a narrower perspective. Faced with the long-standing decline in the United Kingdom's performance (and apparently ignorant of the very recent improvement of the terms of trade caused chiefly by the declining value of the pound sterling), the organization comes to the conclusion that unemployment in Britain (expected to exceed 2.5 million this week) will go on increasing for the next ten years at least. The explanation of this odd state of affairs is simple; the British economy will be managed (by economists) as well as may be expected, but the replacement of manual jobs by machines will just keep ahead of the employment of the unemployed in the service industries. The dice, the forecasters say, are loaded against us.

The difference of style between these two gloomy visions of the future is instructive and surprising. The economists, fond as they are of claiming sensitivity to human behaviour, have blundered by pretending to calculate the difference between two large numbers — the number of people displaced from existing jobs plus that of the nett number of those recruited to the labour force — and the number taken up into new industries with an accuracy of less than one per cent. The conclusion that the problem must be made to go away by import restrictions is familiar but irrelevant. The Council for Science and Society is more perceptive, recognizing that it cannot foretell the future with any accuracy, but that the future is likely to be as painful as was the past. That both the future and the present may be pleasurable is hardly mentioned, but that is now the fashion.

Fusion researchers face fight for funds

Weapons lobby will oppose Reagan's cuts

Washington

Research into inertial confinement fusion (ICF) — the use of lasers or particle beams to ignite pellets of tritium and deuterium — is fighting for its political life. Supporters in Congress are trying to head off a direct frontal attack by the Reagan Administration, which wants to reduce the ICF operating budget by 25 per cent next year. And already it seems virtually certain that funds will dry up for what some see as the most technically-promising way of obtaining usable energy from such fusion reactions, the use of heavy ion beam drives.

ICF research has been a constant battleground for several years. At present the research is funded as a defence-related programme within the Department of Energy, since the explosion of the pellets can be used to study the simulated effects of nuclear weapons. The Armed Services Committee in the House of Representatives has been a particularly enthusiastic supporter, championing the construction of facilities at the Department of Energy's two weapons research laboratories and the Sandia Laboratories in Albuquerque, New Mexico, at a total cost of more than \$1,000 million since 1977.

However, scepticism within the Office of Management and Budget over whether this level of expenditure is justified has recently been fanned by several factors. One is the continuing difficulty being experienced by the laboratories in approaching the point of thermonuclear ignition, and another is the extreme tight-fistedness of the new Administration, seeking in particular to find ways of reducing federal support for applied energy research.

Things were already turning sour before Mr Reagan came into office. When Mr Carter presented his budget for 1982 to Congress in January, he proposed a total of \$219 million to support ICF, the same level of funding as last year with no increase to allow for inflation, and considerably less than the Department of Energy had requested. The Reagan Administration reduced this still further to \$180 million. In particular, it suggested that operating funds for the ICF programme be reduced from Mr Carter's proposed \$140 million to \$106 million, a figure which could result in the lay-off of 300 to 400 scientific and technical staff according to the director of the Office of Inertial Fusion, Dr Richard L. Schriever.

These budget reductions are being sharply contested by the House Armed Services

Committee, which is responsible for overseeing the ICF programme. The committee voted to restore all of the money which the Reagan Administration wants to cut, increasing the department's request for operating funds from \$106 to \$156 million. This represents the single largest increase recommended in the weapons research budget, and was accepted by the full House of Representatives when it voted on the authorizing bill last week. It now passes to the Senate, and then on to the appropriations committees, where it could get a rougher ride.

The Office of Management and Budget is unlikely to give up the fight. There are rumours of demands for even bigger cuts in the ICF budget for 1983. In particular,

officials are keeping up their efforts to quash plans supported by the Armed Services Committee to upgrade the NOVA glass laser facility under construction at the University of California's Lawrence Livermore Laboratory.

Scientists at Livermore strongly refute charges that NOVA represents unnecessarily extravagant expenditure. They point to recent research results which have raised hopes that thermonuclear ignition can be reached on schedule by 1986–87. In particular last year's discovery that much higher efficiencies in beam-target coupling can be achieved using shorter wavelength light. This is already being used as the basis for a planned experiment, using two of

Fredrickson quits without warning

Washington

To the consternation of the US biomedical research community, Dr Donald Fredrickson announced unexpectedly last Friday that he will be leaving his position as director of the National Institutes of Health (NIH) on 1 July.

Dr Fredrickson cites "personal" reasons for leaving the world's largest biomedical research organization, referring in particular to the weight of administrative responsibilities and a desire to return to his research in lipoproteins and lipid transport. He has denied any differences of opinion with the new Administration over the funding of biomedical research, although the speed of his departure has inevitably provoked rumours of problems which have yet to surface publicly.

One point of particular frustration to Dr Fredrickson is said to have been the difficulty of making appointments to top positions at NIH at a time when any such decision requires a special exemption from the current freeze in federal hiring. Several senior administrators have recently left NIH, including Dr Robert L. Goldberger, deputy director for science, who has been appointed provost and vice-president for health sciences at Columbia University in New York.

A respected scientist and a popular administrator, Dr Fredrickson first joined NIH 28 years ago, and was appointed their director by President Gerald Ford in 1975. He had previously been director of the National Heart and Lung Institute, and in 1974 became president of the Institute of Medicine at the National Academy of Sciences.

Dr Fredrickson is said to have a good working relationship with the new Secretary of Health and Human Services, Mr Richard Schweiker. Mr Schweiker had previously been a keen supporter of NIH as a member of the Senate's Human Resources Committee, and is said to have recently agreed to reconfirm Dr Fredrickson's appointment — just as Mr

Joseph Califano had done when Mr Carter became president in 1977.

If Dr Fredrickson is having problems with other members of the new Administration, then they are being well concealed. The conviction that there must be some unexplained reasons for his rapid departure seems partly an expression of disappointment that a popular director is leaving NIH in the hands of an unknown



No whisper where next

successor just when possible budget restrictions could mean a new period of uncertainty. During Dr Fredrickson's tenure at NIH the research budget has increased from \$1,880 to about \$3,500 million, although the rate of expansion has lately slowed considerably.

Last week, in delivering an emotional address to NIH staff at the end of which he announced that he had submitted his resignation to President Reagan, Dr Fredrickson said that NIH were "healthy, strong and in the prime of life", while warning that the essence of the institutes' greatness was "fragile and could be destroyed by careless trustees". He has made no announcement about future plans, apart from telling senior staff that he will spend some months with the National Academy of Sciences. **David Dickson**

NOVA's beams and equipment from the Argus and Siva lasers — an experiment known as NOVETTE — operating with green light. This experiment will make it possible to test parts of NOVA before the full facility becomes operational, a cost-reducing exercise which, it is hoped, will please the sceptics at the Office of Management and Budget.

Contingency plans are also being made at the Los Alamos Scientific Laboratory, where the 35–40 kilojoule carbon dioxide laser ANTARES is due to be completed in 1983, and at Sandia Laboratories, where construction of the first phase of a new particle beam fusion accelerator has just been completed. Each of these programmes would be severely affected if the full cuts proposed by the Reagan Administration go through.

Given the political strength of the Armed Services Committee, and the general enthusiasm for military-sponsored research, it seems unlikely that the full cuts will, in fact, materialize. There is much enthusiasm for the military use of lasers among conservative Republicans, and the White House's National Security Council is studying a proposal for substantial increases in funding for a space-based, high-energy laser system, and has given support to a decision by the Senate to add \$50 million for research in this area.

A more likely casualty of the budget stringencies, however, is heavy ion beam research. From a purely technical point of view, many experts believe that the use of heavy ions to explode (or more accurately, implode) the pellets — as opposed to the light ions being studied at Sandia Labs — present engineering difficulties in designing a reactor. But heavy ions continue to suffer from two major drawbacks.

First, it would cost many hundreds of millions of dollars to build an experimental machine, expenditure that stands little chance of gaining federal support in the prevailing financial climate. And second, unlike either lasers or light ions, such technology has no potential military uses.

Support from the Armed Services Committee has ranged from lukewarm to frigid. And given its own priority of glass lasers, it has given heavy ions a firm thumbs down, not only supporting the Administration's decision to withdraw funding for this area of research (which received \$5 million in the current year), but adding explicit language in the authorizing bill forbidding the department to provide support from other funds.

Supporters of heavy ions had hoped that their cause would be strengthened by a review of ICF previously announced by the Department of Energy's Research Advisory Board (*Nature* 287, 573; 1980). But with the change in administration this review has been put on ice; and with both Congress and the Department of Energy now stressing that this principal interest in ICF lies in its military applications, the prospects for heavy ion drivers look increasingly bleak.

David Dickson

European space success

Ariane takes off

Kourou, French Guiana

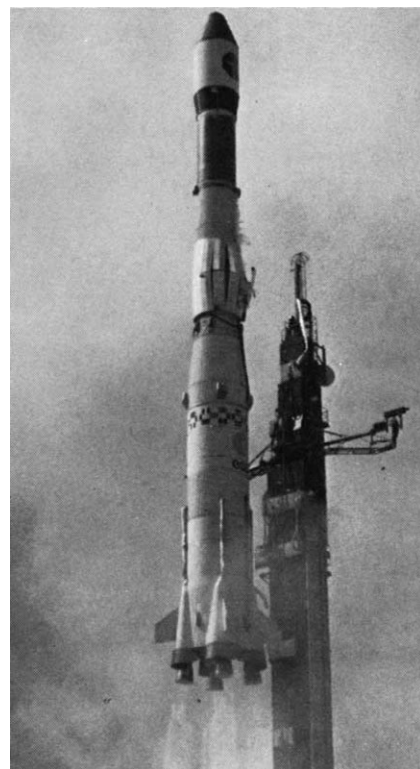
Champagne corks were popping here last Friday after the near perfect third test launch of Ariane, Europe's hope for an independent satellite launcher. During the final countdown, tension was high amongst the largely French team of engineers who had worked to correct the fault in the first stage Viking engines that had led to the failure of the second test launch a year ago. They, at least, are now reassured that the earlier problems have been ironed out.

Lift off was at 12.33 GMT, 19 June, on the third countdown, about 1 hour 10 minutes later than scheduled. The first two attempts were abandoned after a marginal overload in the power supply to the third stage, and when two local radar stations failed to synchronize. In the event, the voltage tolerance was increased and only one radar station tracked the rocket during the first few minutes of flight. Preliminary analysis of data indicated that the trajectory and timing were on target and that the three payload satellites, Meteosat II, Apple — the Indian experimental communications satellite — and a technological capsule, were injected into a suitable geostationary transfer orbit.

Over the next few weeks, detailed analysis of flight data will reveal precisely how the first stage engines performed. The worry had been that high frequency combustion instabilities at 2,300 Hz and 2,700 Hz, which had resulted in the failure of the second launch, might recur. Although the mechanics of the instabilities, not uncommon in liquid fuel rocket engines, is not understood, the problem seems to have been solved largely by redesigning and increasing the size of the orifices in the hydrazine and nitrogen tetroxide fuel injectors.

With two successful test launches behind it, Ariane is now qualified for commercial operation, according to the European Space Agency (ESA). After the fourth and final development launch, scheduled to put ESA's maritime satellite, Marecs, into orbit on 10 November, commercial operation will be handed over to Arianespace, a consortium of European companies. The fourth development flight, which will launch within a particularly narrow window at night, will be a more stringent test of Ariane's capabilities than last week's exercise. Subsequent commercial launches will confirm whether the launcher achieves its design reliability of 95 per cent.

Confidence, however, is now such that work began earlier this week on building a second launch pad at Kourou, finance for which is expected to be approved by the ESA member states later this month. ESA is hoping that last week's success will prompt enough orders to fill the ten launch



per year capacity of the two pads. So far seven firm orders and fourteen reservations have been made, mainly for placing communications satellites into geostationary transfer orbit. The hope is that two or three more firm orders will be placed by the autumn and that several customers, including the Arab countries, which will want to launch two Arabsat telecommunications satellites in 1983 and 1984, and the Australians, will now be prompted to opt for Ariane rather than the United States space shuttle, which, despite its successful inaugural flight, still remains to be qualified.

Meanwhile, ESA's role in Ariane's development is not yet over. Work is under way on Ariane 2 and 3 with payload capacities of 2,000 kg and 2,400 kg respectively, compared with the present capacity of 1,700 kg. Hopes are now high within the Centre National d'Etudes Spatiales (CNES), the French space agency, that the ESA member states will be sympathetic to Ariane 4 with 3,500 kg or payload capacity. CNES is also planning to propose a partly reusable Ariane 5 later this year.

The success of the French ideas, however, will depend on how the other member states decide that ESA's budget should be spent over the next five years. Negotiations so far have ended in deadlock with the three largest contributors, France, Germany and Britain, wanting to go their own ways. But it is not surprising that after contributing more than 60 per cent of the money and effort spent on Ariane development, France has claimed the lion's share of last week's success and is particularly anxious that the programme should not stop there.

Judy Redfearn

Soviet space research

Salyut era ends

Despite news of last Friday's docking of the unmanned Kosmos-1267 probe with the now-empty Salyut-6 space station, the Soviet media are now celebrating the end of the Salyut-6/Soyuz programme and Mr Brezhnev has outlined the next stage of the Soviet space effort, permanent orbital stations served by relays of crews.

Mr Brezhnev was speaking at an Order of Lenin investiture ceremony for cosmonauts Kovalenok and Savinykh, just returned from Salyut-6. The projected permanent space station, which, following Tsiolkovskii, Soviet space planners see as a necessary precursor to manned flights to the Moon and planets, will, said Mr Brezhnev, involve "more complicated tasks" — an indication that that will

probably be a considerable delay before such a station is fully operational. However, the docking of the two unmanned vehicles Kosmos-1267 and Salyut-6 is clearly a step towards working out efficient ways of constructing large orbiting complexes.

The achievements with Salyut-6 have been considerable: 28 cosmonauts visited the station including participants from the Comecon partners, there were 34 manned and unmanned dockings, and about 900 separate biological and medical experiments. Some cosmonauts have remained in zero gravity conditions for as long as six months, and the responses of man to prolonged weightlessness will have been well studied.

Before pressing on with new orbital stations though, there is some unfinished business for Salyut planners. Both France and India have been promised berths for their own cosmonauts on Salyut-type craft, and the Soviet news agency Tass has already reported details of the proposed Franco-Soviet medical programme which is to include a new French *krvotok* (Russian for "blood-flow") device which only needs to be pressed against a cosmonaut's neck to allow the associated computer to record all the necessary details about cerebral blood flow.

Vera Rich

Telescope gets director

Professor Riccardo Giacconi, associate director of the Harvard-Smithsonian Center for Astrophysics, has been appointed the first director of the Space Telescope Institute to be established at Johns Hopkins University in Baltimore, Maryland. The appointment, which takes effect on 1 September, was announced last week by the Association of Universities for Research in Astronomy (AURA), which earlier this year won from the National Aeronautics and Space Administration the contract to manage the new institute.

Professor Giacconi has been an X-ray astronomer since he joined American Science and Engineering in 1958. Over the past few years, at the Harvard-Smithsonian Center for Astrophysics, he has been the scientific and administrative leader for the highly successful Einstein X-ray astronomy satellite. The new institute should be built and staffed by 1982; the launch of the space telescope is scheduled for 1985.

European participation in the project will ensure at least 15 per cent of observing time for European astronomers who may also be represented at the same sort of level among the institute's staff. The European counterpart of the Johns Hopkins institute, the proposed European Co-ordination Facility, is now likely to be sited at Garching, near Munich. That at least is the recommendation put forward by a European Space Agency panel set up to investigate the proposed sites. Alternatives under consideration are the Institute of Space Astrophysics at Frascati, the Royal Observatory at Edinburgh and a joint proposal from the Observatoire de Paris and the Institut d'Astrophysique in Paris. The final decision is expected this week.

Philip Campbell

Endangered species

Belgium joins up

Brussels

The Belgian parliament has now voted to bring Belgium into line with nearly all the other members of the European Community by ratifying the Washington Convention concerning trade in endangered species.

Effectively, this means that at every entry point into the Community, customs officials are now required to ensure that any imports of flora or fauna, or their products, are not listed in the Washington Convention.

The only two EEC countries still to ratify the convention are Luxembourg and the Netherlands. The Netherlands has its own strict legislation and like Luxembourg was only waiting for the other members of the Benelux Union to ratify. The European Commission is trying to have a decision adopted that will allow the Community to speak with one voice.

Legislation is all very well, but there are fears that even now that the last gap has been plugged the demand in Europe for fur skins, alligator handbags, leopard tooth jewellery and the like will ensure a continuation of the trade, for Europe is the biggest market for these products.

Belgium is well known as the European centre of the trade in endangered species. And although imports of animal and reptile skins in 1980 showed a decline for the first time in eight years, they were still worth 234 million Belgian francs (£4

million) according to official figures. Experts at the Belgian section of the World Wildlife Fund think that this is a gross underestimate and that the market value of these skins is likely to be more than ten times as high.

Belgium became the centre for the international trade because, over the past eight years, it has provided the easiest means by which goods could be brought into and distributed around Europe free of further customs controls. The new ruling will, however, require that documentation be presented proving that the animal or animal product in question is required for scientific purposes or for a zoo.

This in itself would be difficult to implement, but the real problem is that most customs officials will not have the knowledge or the motivation to tell, say, an endangered species of canary from an unprotected species and 30,000 vegetable species which it divides into three groups — species in danger of immediate extinction, those threatened with extinction and species in danger in a certain country. Specimens in the first list must be accompanied by both an import and an export permit while those in the second list merely require an export permit.

The first list has become considerably longer since the convention was signed in 1973, and at the Delhi meeting of the 68 member states this year a new proposal was made. Instead of listing those species in danger, the reverse should be done. "Reverse listing" would mean that animal and plant traders would have to prove that their goods are a species listed as safe.

Jasper Becker

UK nuclear power

Marshalling forces

Britain's floundering plans to build a series of pressurized water nuclear reactors are to get a thorough shake-up — with the appointment of Dr Walter Marshall, currently chairman of the United Kingdom Atomic Energy Authority, as coordinator of the programme. The appointment will be formalized this week.

Key technical decisions are falling behind schedule, and although there is no present pressure on electricity supplies the government is said to be worried that the forthcoming pressurized water reactor (PWR) inquiry will fall too close to the next general election, complicating the central economic and social issues with the vexed politics of nuclear power.

Dr Marshall's role will be to ensure that the necessary decisions are taken when they need to be taken, drawing together the confusing interests and bureaucracy of the National Nuclear Corporation (NNC, which will construct the plants), the Nuclear Installations Inspectorate (NII, concerned with safety), and the customers, the Central Electricity Generating Board (CEGB).

The situation has been greatly confused

lately by a number of factors: demands by the Installations Inspectorate that British PWRs be better contained, and more free of radiation on the reactor floor, than their American counterparts; worries in the CEEB that the National Nuclear Corporation might be undercapitalized at its present £10 million; the resignation of the director of the corporation (the post remains vacant); and the publication of parliamentary reports on the government's nuclear programme and the CEEB, both of which were critical of the way in which the need for the programme had been assessed.

Marshall has apparently begun by asking the several interested parties to set down on paper a list of the decisions they consider need still to be made, noting those on which everybody has already agreed. He plans then to draw up a timetable for decision on the outstanding questions, confining his own role to making sure that the timetable will be met. Although the Atomic Energy Authority has a substantial programme of research on the safety of the PWR, Marshall plans to be a judicial chairman, not allowing his own organization as such to have the last word.

Dr Marshall, it is hoped, will fill the managerial vacuum without direct responsibility for the technical decisions themselves, which will remain in the hands of NNC, NNI and CEEB. However it seems inevitable that Marshall, with his forceful and argumentative personality, will stamp his mark on the project.

Robert Walgate

Basel Immunology Institute

Dancing on

Basel

The Basel Institute of Immunology, internationally distinguished for its contributions to fundamental research, last week celebrated its tenth birthday with a series of general scientific talks intended to be intelligible to a wide audience, and a ballet intelligible only to immunologists.

The celebration came at a critical point in the life of the institute, whose directorship passed last summer from its founder, Niels Kaj Jerne, to Fritz Melchers, one of its earliest members (see *Nature* 286, 3; 1980). This has inevitably raised the question of how far the success of the past ten years has been due to the personality of the founding director, expressed both in the organization of the institute and in its intellectual tone. Apart from about ten permanent members, most of whom have been there from its inception, the institute has no tenured personnel: the remaining 50 or so scientists hold two- or three-year contracts and usually leave at the end of them. During their term at the institute, all scientists have strictly equal status.

The other conspicuous feature of the institute under Jerne was the distance kept between the research staff and that of

Hoffman La Roche, the Basel-based pharmaceutical company that supports it. Melchers promised when he took over that the distance would decrease, and indeed Roche laboratory personnel do now visit the institute for discussions with its members. So far, however, much the most effective link between the institute and the company has been Theo Staehelin, now working for Roche, but still temperamentally part of the institute.

A more difficult question is posed by the imminent departure of Susumu Tonegawa for the Massachusetts Institute of Technology after ten years in Basel. In 1976 he was the first to announce the sequence of genomic DNA coding for an antibody molecule and has since been consistently among the leaders in the analysis of cloned immunoglobulin DNA. The question of his replacement is problematic in several ways. Research on cloned DNA is now so brutally competitive, and its practitioners so much in demand, that the loss of Tonegawa may also lose the Basel Institute its lead in the molecular biology of antibody genes. But an attempt to maintain such a lead might in any case have dangers for the institute. It would absorb a substantial proportion of the support provided by Hoffman La Roche, and would almost certainly entail the contravention of the institute's principle of equality and autonomy of all its members.

Melchers, however, recognizes that to abandon research on recombinant DNA altogether is out of the question. Although most of Tonegawa's collaborators are leaving either with him or for other academic laboratories, Hiroshi Sakano remains for about a year and will be joined by one of Tonegawa's earlier collaborators. Melchers is also seeking another molecular biologist who will combine expertise in recombinant DNA technology with a feeling for immunology. Immunoglobulin, he points out, is not the only molecule synthesized by lymphocytes — they also produce important intracellular signalling molecules about which almost nothing is known. Research on membrane proteins is another line Melchers would like to see at the institute.

This emphasis partly reflects Dr Melchers' own research interests, which have been in the clarification of the cellular interactions that trigger antibody production by specialized plasma cells, the subject of one of the three scientific sessions the institute anniversary celebration. The other two dealt with immunoglobulin and the major histocompatibility antigens. One topic conspicuous by its absence was the "idiotypic network" proposed by Jerne to explain the regulation of the immune response and which has long been among the research preoccupations of the institute. Jerne, summing up the meeting, commented only mildly on the omission.

The omission was made good in the ballet, performed with astounding verve at

the end of the celebration by respected scientific members of the institute and their technical staff, ingeniously costumed as a variety of lymphocytes and associated cells.

Particularly notable were performances from Harald von Boehmer, benignly engulfing as a macrophage; Kirsten Fischer-Lindahl, chilling as the male-specific T killer cell, Ivan Lefkovits clinically efficient as a T helper cell, and, at a more molecular level, Charles Wood as a Samurai recombinase cleaving a *corps* of nucleotides that athletically recombined to code for antibodies too numerous to mention.

Miranda Robertson

German nuclear reprocessing

Decision on site

Attempts are under way to break the stalemate in West German nuclear politics. DWK, the German association for reprocessing spent nuclear fuel, has now announced its choice of a site for a small reprocessing plant on the borders of North Hesse and North-Rhine Westphalia. And the federal ministers for research and the interior expressed the view that there are no technical obstacles to a high-level waste storage facility in the salt domes of Gorleben in Lower Saxony.

Those statements are significant because, if followed by the appropriate action, they could ease the burden on utilities applying for licences to build new nuclear plant to satisfy the government that they can safely dispose of nuclear waste. Since 1976, various methods of waste disposal have been a condition of construction and commissioning licences. Hopes that those conditions could be met through reprocessing and final waste disposal in a geological formations were dampened in 1979 when planning permission for a major facility at Gorleben was refused after a lengthy public inquiry.

Since then, storage tanks at power stations have been upgraded to store for longer and permission has been sought to build intermediate storage tanks — which the inquiry had not ruled out — at Gorleben. The search has also been on for alternative sites for reprocessing plants in states whose governments are known to be sympathetic to nuclear power.

So far, DWK has made investigations in two such states: North Hesse and Rhineland-Phalz. The Rhineland-Phalz negotiations are still at an early stage. The idea is to build two small reprocessing plants each capable of 300–350 tonnes throughput a year. A third plant, to be built later at an as yet unspecified site, would have a 700-tonne capacity bringing the total capacity up to the 1,400-tonne originally planned for Gorleben.

DWK's choice of a specific site in North Hesse means that formal application can now go ahead. But permission will not be granted easily. Although the state government under the leadership of Herr

Holger Boerner is sympathetic, opposition within his ruling Social Democratic Party could scupper the scheme. Last weekend, however, Herr Boerner went to his local party for a vote of confidence in his leadership. His mandate as premier of North Hesse was supported, but opposition to the reprocessing plant together with a plan to extend Frankfurt airport was voiced. The likely outcome now is that the reprocessing plant will go-ahead but local opinion will be firmly divided.

Meanwhile, the local inhabitants of Gorleben have reopened discussion on whether a high-level waste storage facility should be built in the area. At a public meeting in May, several objections to the technical feasibility of the facility were aired: the salt domes are not covered by a layer of clay, they are in contact with ground water in some places, the region may not be as geologically stable as originally thought and natural gas is extracted from the eastern section of the domes in East Germany. Nevertheless, the federal government seems convinced that the Gorleben salt domes offer the best site within West Germany and has said that, even if conditions are not ideal, problems can be overcome by tailoring the waste to suit the geological formation. The government is still committed to an operational storage facility by the 1990s and Gorleben still looks like the best option.

Judy Redfearn

European environment

Win some, lose some

Luxembourg

Environment ministers of the European Community are in the throes of attempting to draw up regulations on the discharge of noxious substances, outlined in the 1976 directive ENV 131. On 11 June, they again came within a hair's breadth of agreement on the seemingly intractable problem of common standards for the discharge of mercury into the aquatic environment. But one outstanding issue remains.

The United Kingdom is continuing to defend its right to apply qualitative objectives for the discharge of mercury from both old and new plants, while Italy and France insist that limiting values should be obligatory for assessing pollution from new plants. The British, it is maintained, are defending a point of principle while their continental partners are protecting their commercial interests. Nevertheless, British officials point out that the Italians, for instance, would find it hard to enforce the proposed maximum values. It is hoped that the disagreement can be resolved within the next three months.

There was further agonizing about the council's decision to postpone the adoption of the so-called Seveso directive on the prevention of major industrial accidents. The French have held up agreement, pleading that the new government

has not had time to adopt a fresh policy, but hinting that the Mitterrand regime will be more sympathetic than its predecessor to the outstanding problem of trans-frontier notification.

Another subject for debate is the threat to the ozone layer posed by chlorofluorocarbons (CFCs 11 and 12). At the end of this year, the European Community is due to review its commitment to reduce the use of these materials in aerosols to 30 per cent of 1976 levels. The commission's report is unlikely to satisfy either side in the dispute.

Although sales of CFCs 11 and 12 as aerosol propellants fell in 1980 to 28.2 per cent of 1976 levels and are expected to drop to 35 per cent this year, sales for other uses such as refrigeration and plastic foam have risen by 34.7 per cent. The commission's report implies that the situation is well in hand, but favours reducing the production of CFCs for uses other than in aerosols. Negotiations are also to be pursued within the framework of the United Nations Environmental Programme and the Overseas Environmental Council for Development. The council seemed content to assume that, even if the ozone layer is being destroyed at an annual rate of 5–7 per cent current measures are adequate. Even so, the Community appears to be taking a daring step; there is nothing like a consensus among its members that the present restrictions are necessary.

Surprisingly, the United Kingdom is pushing for stringent Community restrictions on lead pollution, on the lines recently adopted in Britain, which would mean a lowering of the lead level in petrol from 0.4 to 0.15 grammes per litre. West Germany and Denmark will be enforcing this limit, and it now seems likely that the Commission's 1976 proposed directive on the permissible level of lead in air will also be taken off the shelf. The council has agreed in principle to endorse an environmental standard of 0.2 milligrammes per cubic metre of air, and has directed that the developer should shoulder the costs of the assessment, that the information should be publicly available and that there should be consultations with neighbouring states. However, 130 reservations were tabled against the directive at the beginning of the council meeting.

With an air of embarrassed pride, the ministers also announced agreement on two measures designed to prevent further disasters similar to that of the *Amoco Cadiz* — drawing up a compendium of the properties of hydrocarbons and a list of oil-fighting equipment in the Community. The more important proposal to keep a register of oil tankers most at risk was resisted by the Greeks and passed on for consideration by the transport ministers. The two proposals agreed are the only measures produced since the European Community announced its crusade against oil tanker disasters following the 1978 *Amoco Cadiz* shipwreck.

Jasper Becker

Water works for the Cuna villagers

Panama

For the Cuna indians of Panama progress still means hard physical work. But as a result, the inhabitants of the whole Cuna village can now enjoy clean running water.

The Cunas are colourful tribes of interbred indians who live in palm and bamboo huts on the beautiful San Blas islands — a cluster of over 360 small coral islands off Panama's Caribbean coast. The location allows an almost completely autonomous existence.

The island village of Ticantiki, with its 900 inhabitants, is situated on Devil Cays, only 600 yards from the mainland. It is this proximity to the mainland that has allowed fresh water to be piped under the sea to the island.

A new dam has recently been built up in the hills of the mainland as part of a government project involving the labour of the local indians. In indian terms, the dam is three hours' walk through the jungle from the coast. From the dam, fresh water flows through plastic piping down to the island at a rate of 60 gallons per minute.

The building of a water filter has been left entirely to the indians. The filter is situated a half-hour's walk from the coast, about 2 miles. The shell of the filter

is made of reinforced concrete and comprises two bins, each about 8 ft square and 6 ft deep. Carefully graded stones are being used to fill the bins and provide the filter surface. Water is sprinkled from overhead through holes in the plastic piping. The filter, though simple in concept, is effective. When it is finished the indians hope to be able to increase the water flow to 80–90 gallons per minute.

But one may ask, where in a tropical forest do you find enough stones for a filter of this size? The answer is, at the mouth of the river. Here, children help in the work by grading stones. The men of the village carry the stones up to the filter in sacks suspended from a pole on their shoulder. This load weighs about 50 kilos, yet these tough, stocky indians stride through the jungle as though it were no more than a sack of cotton.

The congress of the village dictates that any able-bodied man who does not carry up a load should pay \$5 a day to the community. For a culture that lives by fishing, tending plantations and, increasingly tourism, this is a large sum of money.

And the next project for this ambitious little village? To pipe their fresh water to every household.

Patricia Dent

CORRESPONDENCE

Huxley on *Nature*

SIR — You quote the president of the Royal Society (*Nature* 4 June, p.373), speaking at the centenary celebrations for the British Museum (Natural History), as criticizing *Nature* for having "printed more than 30 letters . . . relating to cladistics, which would have been unintelligible to 99 per cent of the readership".

But, sir, would that not apply equally to many of the other topics dealt with, and rightly dealt with, in the pages of your esteemed journal? Elitist it may be, but it is to be hoped that you will continue to give space to technical scientific matters, especially when these are a subject of controversy, commensurate with their interest and importance, even though this may be apparent to no more than a small minority of your readership.

C. B. GOODHART

University Museum of Zoology,
Cambridge, UK

Cabal of referees

SIR — The article in *Nature* referring to the policing of the publication of scientific literature (9 April, p.433) raises a major issue.

We suggest that the main reason for the problems now surfacing and referred to in this article is the existence of a refereeing system which has turned into a cabal and like all secret societies is plagued by corruption.

This problem has already been raised by an assistant editor of the journal *Nuclear Physics*, who felt that the present anonymous system of peer review results in injustices to authors and is inappropriate to the spirit of science (P. Robertson *New Scientist* 71, 410; 1976.)

Quite clearly, the reputation of this system is now at stake. Could it be that the scientific community at large is not yet mature enough to handle the problems that may arise from a more open system?

M.A. GILLMAN
F.J. LICHTIGFELD

Medical School,
University of the Witwatersrand,
Johannesburg, South Africa

Book for safety

SIR — In an editorial on nuclear politics (*Nature* 4 June, p.368) you suggest that governments should be responsible for safety and responsive to the views of the electors. This is to recognize the democratic imperative, but you might have added the need to inform the electorate about safety matters. With these, as with all difficulties, the strength of a democracy depends on the sense of its citizens.

One of our functions at the National Radiological Protection Board is indeed to provide information on radiation safety, and we do so in several ways. Because of public anxiety about nuclear issues, we discuss them impartially in a new general publication *Living with Radiation* (Her Majesty's Stationery Office, London, 1981; 50 pence) which we hope will promote cool debate and calm judgement.

M.C. O'RIORDAN

National Radiological Protection Board,
Chilton, Didcot, UK

Wheel re-invented

SIR — The recent report by K.S. Jayaraman (*Nature* 28 May, p.277) describing the "gunny-bag air conditioner" developed by the Central Building Research Institute in India, provides an excellent example of the redevelopment of wheels at government expense. The process of passive cooling by roof-surface evaporation is as Jayaraman says suitable not only for India and other developing countries in the tropics but for large areas of American suburbia where for many years the advent of summer has been marked by the temporary placement of lawn sprinklers on the ridge of the roof.

J.P. BENTLEY

University of Oregon Health Sciences Center,
Portland, Oregon, USA

Doomed subspecies

SIR — How would a sociobiologist explain the existence of anti-sociobiologists?

Undoubtedly, most people in the world would claim to be basically generous, caring, and unselfish individuals. Sociobiology, with its fundamental emphasis upon the selfishness of human nature, thus represents a minority view.

But even if the majority view of human nature is untrue, being the result of genes which have been selected for such "deception", might anti-sociobiologists nevertheless receive greater fitness benefits from this majority?

Might one expect to see the demise of sociobiology over evolutionary time due to the decreased fitness of sociobiologists relative to anti-sociobiologists?

ARTHUR MELTZER

Syracuse, New York, USA

Divergent usage

SIR — I would like to note a common and unfortunate misuse of terms. In his *Origin of the Species*, Darwin noted that similarities among organisms could arise from *convergence* because separate lines of descent independently acquired a function, or by *divergence* from a common ancestor. I do not know who coined the words, but Darwin used *analogy* to describe the former and *homology* to describe the latter. That useful distinction has served biologists well for over a hundred years.

The clarity of these concepts is now being widely abandoned as illustrated by your associate editor who presents the conclusion "that the homology is due to divergent rather than convergent evolution"¹. She thus presents us with a choice between a tautology (homology due to divergence) and a contradiction in terms (homology due to convergence).

One should use *similarity* for the meaning that homology must have in the above quote for reasons far more important than a personal preference for Latin over Greek-derived words or preserving a time-honoured usage. Similarity is an *observation* while homology is an *inference* about its origins. To use homology with both meanings is to confuse the data with the hypothesis. And how much simpler it is to conclude, "The similarity is homologous rather than analogous".

WALTER M. FITCH
RONALD L. NIECE

Department of Physiological Chemistry,
University of Wisconsin-Madison,
Madison, Wisconsin, USA

1. Robertson, M. *Nature* 290, 187 (1981).

Credit due to Nabi

SIR — It has recently been suggested in the columns of *Nature* that I am the mysterious Isidore Nabi (*Nature* 23 April, p.623). I would like to do what I can to clarify the situation. Let me state categorically that any assertion that Isidore Nabi is none other than R. C. Lewontin is incorrect. Let me offer a few corroborative details.

(1) According to his biography in *American Men and Women of Science*, Dr Nabi is 71 years old, received his bachelor's degree from Cochabamba University, and, among other things, has lectured and carried out research at the University of Venezuela for five years. I, on the other hand, am 52, have never even heard of Cochabamba University, and have never been south of Mexico City.

(2) Dr Nabi is the editor of the journal *Evolutionary Theory* on whose editorial board I also appear by name, and I also find him listed as a member of the Evolution Society of which I once had the honour of being president. Why would Professor Van Valen, managing editor of *Evolutionary Theory*, list me on the editorial board if I were also, under a different name, editor of that worthy journal? And what in the world would I do with an extra copy of *American*

breeding, 02. prof exp. res basic plant genetics, dept. union breeding, ministry agric. Egypt. 65-67; res dir plant breeding. Dunn Seed Co. 68-71; prof biol. Odessa Col. 70-75; MEM FAC SCI-MIDLAND COL. 75; CHMAN DEPT SCI STUDIES. 77; Mem. Am Genetic Assn; Am Soc Agron. Res. Industrial research on development and improvement of cotton, the inheritance of quantitative and economic characters in field crops, with major emphasis on disease resistance, insect tolerance and fiber qualities in upland cotton. Mailing Add: Dept of Sci. Midland Col. Midland TX 79701.

NABI, ISIDORE, b Brno, Czech, July 22, 10; m 30; c 6. POPULATION BIOLOGY. Educ: Cochabamba Univ. AB. 30; Nat Univ Mex. MD. 36. Hon Degree: PhD. Cochabamba Univ. 50; LLB. Nat Univ Mex. 39. Prof Exp. Petrol geologist. Ministerio de Fomento, Venezuela, 40-42; instr. biol. Hunter Col. 43-47; resident path. Kings County Hosp. Brooklyn. 47-49; ed & publisher. Boletín de Medicina Forense. Caracas. 49-51; lectr & res assoc path. Univ Venezuela. 51-56; Guggenheim fel biol. Yeshiva Univ. 56-57; res assoc pharmacol. NY Univ. 62-65; res assoc anat. 65-67; evolutionary biol. 67-71. RES ASSOC BIOL. UNIV CHICAGO. 71; Concurrent Pos. Consult. Standard Oil Co. 43-47 & Kings County Coroner. 47-49; NIH res grant. 65. Mem Soc Study Evolution. Am Col Legal Med. Int Acad Path. Res. Cytopathology; forensic cytology; paleocytopathology. Mailing Add: Dept of Biol Univ of Chicago Chicago IL 60637

NABIGHIAN, MISAC N. b Bucharest, Rumania, Dec 5, 31; US citizen; m 66; c 2. GEOPHYSICS. Educ. Mining Inst Bucharest. BS. 54; Columbia Univ. PhD (Geophysics). 67. Prof Exp. Geophysicist. Geol. Coun. Bucharest. 55-57; res scientist. Geophysics Inst. Bucharest. 57-62; res assoc. 62-65; res assoc. 65-67. GEOPHYSICIST

Naturalist when I hardly know what to do with my own?

(3) Isidore Nabi is the author of several important works which, I am sorry to say, are not at all of my creation. I refer in particular to his brilliant "On the Properties of Motion" which is as yet unpublished but widely circulated and known, and his seminal work "An Evolutionary Interpretation of the English Sonnet".

I have recently received a letter from Professor Van Valen saying that he has been identified as the Isidore Nabi who wrote the letter to *Nature*, an assertion which he denies. Thus, confusion multiplies. I hope that this letter has thrown some light on the situation.

RICHARD C. LEWONTIN
Museum of Comparative Zoology,
Harvard University, Massachusetts, USA

NEWS AND VIEWS

Cosmic quadrupole?

from Richard A. Muller

"THE truly marvelous thing about science is the great return in theory that one obtains for such a meager investment in fact." — Mark Twain.

In no other field of science is Mark Twain's perception as true as in cosmology. There are so few 'facts' that constrain cosmology that the discovery of a new one is a major event. One of the most recent such discoveries is the anisotropy (variation in intensity with direction) in the cosmic microwave background radiation. It is now well established^{1,2} that the intensity of the radiation varies by about 0.2 per cent (6 mK of the 3K signal), with maximum intensity towards the Virgo cluster of galaxies, and minimum in the opposite direction. The intensity of the modulation varies smoothly between these extremes as the cosine of the angle in the sky, suggesting that the anisotropy is due to a Doppler shift from motion of the Milky Way galaxy with respect to the radiation. This cosine modulation has recently been observed^{3,4} to have a blackbody spectrum in the frequency range from 19 to over 100 GHz. This observation has eliminated lingering fears that the anisotropy might be a spurious effect due to interference by galactic synchrotron emission.

Two groups are now reporting a new 'fact' — a second harmonic in the cosmic background anisotropy. R. Fabbri, I. Guidi, F. Melchiorri and V. Natale³ in Italy have observed a second harmonic with amplitude 0.9 (+0.4, -0.2) mK and axis consistent with that of the first harmonic. S. Boughn, E. Cheng and D. Wilkinson⁴ at Princeton report a 'quadrupole moment' Q_2 (that is, a second harmonic proportional to $\cos^2\delta \cos 2\alpha$ where δ is declination and α is right ascension) of magnitude -0.54 (+ -0.14) mK. Observations of second and higher order harmonics are particularly exciting because, unlike the cosine anisotropy, they are not easily explained by local effects such as proper motion. Rather they give us information about the state of the Universe at very early times.

The existence of the cosmic microwave radiation was predicted by G. Gamow, R. Alpher and R. Hermann⁵, and independently by R. Dicke, P. Peebles, P. Roll and D. Wilkinson⁶ before the radiation was discovered by A. Penzias and R. Wilson⁷ in 1965. Their theories have formed the basis of what is now referred to as the 'standard model' of the early Universe. In this model, the microwave radiation was once thermal radiation in equilibrium with the hot plasma that filled all space. About a half million years after the big bang, expansion of the Universe caused the previously opaque plasma to cool sufficiently to become transparent to the radiation, and since that time the signal has been travelling essentially unimpeded through the Universe. When one observes the radiation, one is literally looking at the shell of matter that last scattered it. The radiation comes from a more distant region of space, and was emitted earlier in time, than any other observed signal. The microwave radiation literally forms the spatial 'background' in front of which all other astrophysical objects lie.

There are many effects that could give rise to anisotropy in the radiation, including long-wavelength gravitational radiation, an overall rotation to the Universe, an anisotropic Hubble expansion and nonuniformity in the density of the Universe at the time of decoupling. Only the cosine or 'dipole' term (so called because a dipole distribution is proportional to the first spherical harmonic) can come from the Doppler shift due to the velocity of the observer with respect to the distant matter that last scattered the radiation.

Calculations indicate that the reported quadrupole anisotropy is consistent with that expected, based on observed inhomogeneities in the distributions of galaxies, and extrapolations in scale and in

time to distributions at the time of decoupling^{8,9}. In fact theorists were beginning to become uncomfortable with the absence of high harmonics in the anisotropy, since it was difficult to see how the present very lumpy Universe could have evolved from a highly uniform one in the short period of 10 to 20 billion years.

Our sigh of relief may be premature, however, because the experimental status is not totally resolved. Anisotropy measurements are very difficult, for they require measurement of millikelvin signals in the presence of system noise of typically several hundred Kelvin. Galactic emission and earth-shine are highly anisotropic, and difficult to eliminate. The history of anisotropy measurements is not reassuring, for published data have revealed large nonstatistical fluctuations subsequently shown to be due to systematic error. And most importantly, the present experiments do not agree within statistical errors. In 1979 G. Smoot and P. Lubin¹⁰ at Berkeley published a measurement of Q_2 consistent with zero: +0.06 (+ -0.2) mK. This value differs from the Princeton value by several standard deviations. Even the Princeton dipole terms differ significantly from those of Berkeley. No one doubts the reality of the dipole anisotropy, but the disagreement does show that one must be wary of accepting statistical errors as the true errors. The three experiments are very different in design, and the systematics could be very different. The Berkeley experiment sampled regions of the sky in both the Northern and Southern Hemispheres. The Princeton experiment had continuous sky coverage, but only in the northern sky. The Italian experiment had the poorest sky coverage of all; its high frequency makes it insensitive to synchrotron background, but more sensitive to thermal radiation from galactic dust. Since galactic emission can be point-like as well as diffuse, it is impossible to guess which experiment is most sensitive to unknown backgrounds. Smoot and Lubin are cautious, and willing to give as their

Richard A. Muller is Professor of Physics at the University of California, Berkeley and Faculty Senior Scientist at the Lawrence Berkeley Laboratory.

final upper limit 1 mK for a quadrupole term. Fabbri *et al.* are similarly cautious. They consider their effect, 4.5 standard deviations from zero, to be only 'suggestive' of a quadrupole anisotropy. Only the Princeton group is convinced that they have been able to take all systematic errors into account.

If their quadrupole term is verified, then the Princeton group will deserve credit for their unambiguous announcement of its discovery. But I think it is premature to accept the effect as proven. Caution is called for, in part because the experiments are difficult, in part because of the discrepancies between the existing

experiments, and in part because the results are close to those expected. Fortunately we may not have long to wait for verification. The Princeton group has already flown in their balloon gondola a maser receiver with considerably improved sensitivity, and the Berkeley group is about to fly a 90 GHz cryogenic system. The frequency behaviour of the quadrupole term will clearly identify it as cosmic background or as galactic interference, and in a few years we should obtain a detailed map of the 3K signal, including higher harmonics, from NASA's Cosmic Background Explorer satellite (COBE). The coverage and sensitivity of COBE are likely to give

cosmologists a whole handful of new facts to inspire and constrain their theories. □

1. Smoot, G.F., Gorenstein, M.V. & Muller, R.A. *Phys. Rev. Lett.* **39**, 898 (1977).
2. Cheng, E.S., Saulson, P.R., Wilkinson, D.T. & Corey, B.E. *Astrophys. J. Lett.* **232**, L139 (1979).
3. Fabbri, R., Guidi, I., Melchiorri, F. & Natale, V. *Phys. Rev. Lett.* **44**, 1563 (1980).
4. Boughn, S.P., Cheng, E.S. & Wilkinson, D.T. *Astrophys. J. Lett.* **243**, L113 (1980).
5. Alpher, R.A., Bethe, H.A. & Gamow, G. *Phys. Rev.* **73**, 803 (1948); Alpher, R.A. & Herman, R.C. *Rev. Mod. Phys.* **22**, 153 (1950) and earlier references therein.
6. Dicke, R.H., Peebles, P.J.E., Roll, P.G. & Wilkinson, D.T. *Astrophys. J.* **142**, 414 (1965).
7. Penzias, A.A. & Wilson, M.L. *Astrophys. J.* **142**, 419 (1965).
8. Peebles, P.J.E. *Astrophys. J. Lett.* **243**, L119 (1981).
9. Silk, J. & Wilson, M.L. *Astrophys. J. Lett.* **244**, L37 (1981).
10. Smoot, G.F. & Lubin, P.M. *Astrophys. J. Lett.* **234**, L83 (1979).

Sexual differentiation of the brain

from Bruce S. McEwen

TESTOSTERONE acts during a limited period of perinatal development in male mammals to influence permanently the development of reproductive behaviour and neuroendocrine function, as well as to promote development of the reproductive tract. Although for some years the brain, and not the pituitary, has been presumed to be a site of testosterone action during development, it is only recently that there has been significant progress in understanding what actually happens to the brain during sexual differentiation.

Neuroanatomical studies have led the way, beginning with the first observation of sex differences in the sizes of nucleoli and cell nuclei in brains of gonadally intact¹ and hormone-treated² adult rodents and primates³. This initial work left open the question of whether hormone action during early development or during adult life was responsible for the morphological features. This question was answered in rat by the electron microscopic studies of Raisman and Field⁴, which revealed that a sex difference in patterns of synaptic organization in the preoptic area of the adult rat depends, at least in part, on exposure to testosterone immediately after birth. Since then, studies at the light and electron microscopic levels have revealed sex differences in central nervous system morphology in songbirds^{5,6} and rodents⁷⁻¹². Among the more notable features of some of these sex differences is their occurrence in neural pathways and brain regions related to sexually dimorphic and hormone-dependent neural events such as song in birds^{5,6}, penile movement¹¹ and gonadotropin secretion¹⁰ in rats.

How do these sex differences arise during the critical period of early perinatal neural development in mammals and birds? One mechanism

being investigated is testosterone-stimulated neurite growth, which has been demonstrated in organ culture¹³ where it appears to emanate from developing neurones that contain hormone receptors¹⁴. According to current ideas regarding the sequence of cellular events in neural development¹⁵, neurite growth stimulated by gonadal hormones might be a sufficient condition for the production of differences in cell number and cell size as well as in the patterns of synaptic connections that are the major features of the sex differences described in the adult brain⁴⁻¹². In spite of the appeal of such a scheme, it would be premature to overlook possible testosterone effects on other aspects of neural development, including the reinitiation of cell division or the promotion of cell death, and the production of substances, like nerve growth factor, which stabilize synaptic contacts. Furthermore, testosterone might alter the neurotransmitter phenotype of developing neurones — a possibility suggested by the actions of glucocorticoids on neurotransmitter type in developing sympathetic ganglia^{16,17}.

Hormonal stimulation of sexual differentiation occurs at a stage of neuronal development when considerable cellular differentiation has already occurred — oestrogen^{18,19} and androgen^{20,21} receptors and enzymes involved in testosterone metabolism to oestradiol²² and 5- α -dihydrotestosterone²³ have already been laid down in certain neural cells. These enable particular neurones to respond to the hormonal signal, and it is known that both androgens and oestrogens arising from testosterone participate in sexual differentiation²⁴, albeit in differing degrees depending on the species²⁵. Oestrogen and androgen receptors are already present in the hypothalamic area of female rodents in the last trimester of gestation^{19,20} from around the time of final cell division²⁶. There are also indications that oestrogen receptor levels in male and female rat brain start to

increase shortly before birth¹⁸ and that androgen receptor levels increase from one week after birth^{21,27}. There are at present no indications for rat or mouse that, at these early ages, sex differences exist in levels of neural oestrogen and androgen receptors. Rather, it is the sex differences in testosterone levels in late fetal life²⁸ resulting in differential occupation of receptor sites in male and female brains^{29,30} which provide the bias leading to sex differences.

Uncovering the factors, hormonal or otherwise, which govern the first appearance and subsequent perinatal increases of oestrogen and androgen receptors and testosterone-metabolizing enzymes in both male and female brains represents a major challenge for our further understanding of the origins of brain sex differences.

1. Pfaff, D.W. *J. Endocr.* **36**, 415 (1966).
2. Dörner, G. & Staudt, J. *Neuroendocrinology* **3**, 136 (1968).
3. Bubenik, G.A. & Brown, G.M. *Experientia* **29**, 619 (1973).
4. Raisman, G. & Field, P.M. *Brain Res.* **54**, 1 (1973).
5. Nottebohm, F. & Arnold, A.P. *Science* **194**, 211 (1976).
6. Gurney, M.E. & Konishi, M. *Science* **208**, 1380 (1980).
7. Greenough, W.T., Carter, C.S., Steerman, C. & DeVoogd, T. *Brain Res.* **126**, 63 (1977).
8. Gorski, R.A., Gordon, J.H., Shryne, J.E. & Southam, A.M. *Brain Res.* **148**, 333 (1978).
9. Reihely, M. *Neuroendocrinology* **28**, 82 (1979).
10. Matsumoto, A. & Arai, Y. *Brain Res.* **190**, 238 (1980).
11. Breedlove, S.M. & Arnold, A.P. *Science* **210**, 564 (1980).
12. Loy, R. & Milner, T.A. *Science* **208**, 1281 (1980).
13. Toran-Allerand, C.D. *Brain Res.* **106**, 407 (1976).
14. Toran-Allerand, C.D., Gerlach, J.L. & McEwen, B.S. *Brain Res.* **184**, 517 (1980).
15. Lund, R.D. *Development and Plasticity of the Brain* (Oxford University Press, New York, 1978).
16. McLennan, I.S., Hill, C.E. & Hendry, I.A. *Nature* **283**, 206 (1980).
17. Jonakait, G.M., Bohn, M.C. & Black, I.B. *Science* **210**, 551 (1980).
18. MacLusky, N.J., Lieberburg, I. & McEwen, B.S. *Brain Res.* **178**, 129 (1979).
19. Vito, C.V. & Fox, T.O. *Science* **204**, 517 (1979).
20. Vito, C.C., Wieland, S.J. & Fox, T.O. *Nature* **282**, 308 (1979).
21. Lieberburg, I., MacLusky, N.J. & McEwen, B.S. *Brain Res.* **196**, 125 (1980).
22. Reddy, V.V.R., Naftolin, F. & Ryan, K.J. *Endocrinology* **94**, 117 (1974).
23. Denef, C., Magnus, C. & McEwen, B.S. *Endocrinology* **94**, 1265 (1974).
24. Goy, R.W. & McEwen, B.S. (eds) *Sexual Differentiation of the Brain* (MIT Press, Cambridge, 1980).
25. Baum, M.J. *Neurosci. Biobehav. Rev.* **3**, 265 (1979).
26. Ifft, J.D. *J. comp. Neurol.* **144**, 193 (1972).
27. Attardi, B. & Ohno, S. *Endocrinology* **99**, 1279 (1976).
28. Weiss, J. & Ward, I.L. *Endocrinology* **106**, 306 (1980).
29. Westley, B.R. & Salaman, D.F. *Nature* **262**, 407 (1976).
30. Lieberburg, I., Krey, L.C. & McEwen, B.S. *Brain Res.* **178**, 207 (1979).

Burce S. McEwen is at the Rockefeller University, New York.

Portuguese workshop on young stars

from a Correspondent

T TAURI stars are cool stars distinguished by spectra which show emission lines resembling those of the solar chromosphere and indicate a surrounding, dense, high-temperature region. They vary in light and are generally thought to be young because they are associated with dust clouds, lie above the main sequence of normal hydrogen-burning stars in the Hertzsprung–Russell diagram and show high abundances of lithium, an element destroyed by nuclear processes at and near the surface early in the main-sequence lifetime of the cool stars. T Tauri stars also show strong continua in X rays, the ultraviolet, infrared and radio which indicate a variety of processes in the circumstellar regions around the stars.

At a recent meeting* on T Tauri stars, one of the main questions discussed was whether the chromospheres contain inflowing or outflowing material. It might be expected that such newly formed stars would still be accreting interstellar material, yet in almost all stars the emission lines show a dip on the blue side. If this dip is due to absorption, it is perhaps natural to attribute it to outflowing material. However, Bertout (University of Heidelberg) and Krautter (European Southern Observatory) pointed out the great attractions of the infall model which provides a region of elevated temperature near the star where the infalling gas is brought to rest. Penston (Royal Greenwich Observatory) noted that all the MgII line profiles observed to date by the International Ultraviolet Explorer Satellite show an absorption or absence of emission on the blue wing and Lamers (University of Utrecht) agreed that in other types of star this is a symptom of mass loss. The most complete model presented for the flow around a T Tauri star was the stellar wind model presented by Lago (University of Oporto) for the star RU Lupi. The spectra of this star provide a series of constraints on a wind model linking the flow velocity (from the line widths) with the density and temperature (from various line ratios). A satisfactory model is obtained if the wind is driven by Alfvénic waves. Schwartz (Queen Mary College, London) remarked that the energy flux carried away in this wind as kinetic energy alone is very high, being some 20 times the luminosity of the Sun.

Observations from space are now making important contributions. Gahm (University of Stockholm) reported that

the ultraviolet spectra of some stars showed evidence of a hot gaseous continuum and emission lines from gas at 100,000 to 200,000 K are seen in most stars. The situation for hotter gas at a million degrees (that is, a corona) was not so clear. A few stars were detected as X-ray sources by the Einstein Observatory but most were not. An absence of X-ray emission might be due to absorption in an overlying wind but attempts had also been made to find emission from coronae in forbidden lines from highly ionized species like Fe^{9+} and Fe^{13+} . These would not be absorbed in the same way and in some stars the limits were so good that one could say that the corona was much weaker relative to the chromosphere than it was in the Sun. A compatible result was presented by Kuhi (Lick Observatory) who showed from his analysis of the X-ray data that there was an anti-correlation of the X-ray luminosity of the detected T Tauri stars with $H\alpha$ equivalent width. In fact the X-ray fluxes, and maybe therefore the coronae, are weaker than for normal stars of the same spectral type. These results suggested some important differences between the winds round T Tauri stars and the solar wind.

Concerning the presence of dust around T Tauri stars, the workshop accepted compelling evidence that, at least for some stars, alternative explanations for the polarization and infrared excesses were no longer tenable. Bastien (University of Bonn) reported that as the polarization was unchanged between $H\alpha$ and the nearby continuum, the polarizing region must be outside the emission line zone and scattering by electrons cannot be the polarizing mechanism. Cohen (NASA — Ames) showed infrared spectra with strong absorption from ice and silicates which demonstrate that solid particles and not electrons are also the source of the infrared emission. In addition, it is difficult to account for the emission observed from aircraft at wavelengths near $100\ \mu\text{m}$ in any other way, although there may still be a role for gaseous emission processes (free-free emission) at radio wavelengths. Panagia (University of Bologna) discussed this point, noting that for T Tauri, the best observed case with an $S_\nu \propto \nu^{-0.7}$ spectrum, the data could be explained by free-free emission in a constant velocity wind. However, in other cases where the radio spectral index is as yet unknown, other models are possible and include the infall model or the origin of the radio emission in 10^9 K flares. Bertout and Cohen presented data indicating that some of the radio sources were variable. The energy necessary to maintain the ionized region can be very large; approximately the same

amount as that emitted at the photosphere.

Lang (Tufts University) discussed the environment from which T Tauri stars are born. They lie in dense dark molecular clouds with densities of order $n_{\text{H}_2} \sim 10^3$ and temperatures near 20 K. The ultraviolet emission spectra also show evidence for molecular hydrogen. Brown (University of Oxford) advanced an explanation in which the strong $\text{Ly}\alpha$ radiation from the star fluorescently excites the molecules in the surrounding medium. Detailed investigation of the mechanism gives the excitation temperature in the molecular cloud and suggests shocks as the heating mechanisms. Cohen also discussed the relationship to T Tauri stars of the Herbig–Haro objects, small nebulae often found in the vicinity of the young stars. He reported that both previously proposed theories of these objects had been found correct in different cases. Some H-H nebulae were ‘mirrors’ scattering the light of the T Tauri stars; others were shock-excited regions. In the latter, the mechanical energy flux required to be emitted from the stars to drive the exciting shocks is very high and Cohen reported that Herbig had in one case found proper motions representing transverse velocities of $300\ \text{km s}^{-1}$. If the wind was spherical the energy carried away was extremely high!

Throughout the workshop, participants were comparing T Tauri stars with other types of star which shared some of their properties. Doazan and Thomas (University of Paris) stressed the similarities with Be stars (hot emission line stars) in which several of the same circumstellar phenomena are seen — indeed one subgroup, the Herbig Ae and Be stars, are almost certainly hotter analogues of the T Tauri stars. However, ‘classical’ Be stars seem to be distinguished by the absence of dust. Felli (Arcetri Observatory) also pointed out the similarity to BN objects. These are infrared stars without optical counterparts in H II regions and are named after the prototype in the Orion Nebula discovered by Becklin and Neugebauer.

Kuhi gave an account of his recent, very thorough programme on the rotational velocity of T Tauri stars. He noted that the oft-reported fact that T Tauri absorption lines are ‘washed out’ could either mean that they were broad or were filled in by emissions. He had avoided some of these types of difficulty by examining the Fourier transform of the absorption spectrum. He confirmed that some T Tauri stars had rotational velocities much higher than those normal for late-type stars but also found that many were rotating too slowly

*A workshop on T Tauri stars was held in mid-April at Ofir near Oporto, Portugal. The meeting was organized by the University of Oporto and jointly sponsored by the Portuguese government, two Portuguese agencies, INIC and JNICT, and the Fundação Gulbenkian.

($\leq 25 \text{ km s}^{-1}$) to be detected. Only those stars, more massive (according to their evolutionary tracks) than about $1.5 M_{\odot}$, were likely to be rapid rotators — a result mirroring what occurs on the main sequence and strongly suggesting that stars have solved their angular momentum problems before the T Tauri phase. His result may suggest that rapid rotation is not the main cause of the T Tauri phenomenon but one should be cautious because the stars with the strongest emission effects do not have photospheric absorption lines enabling their rotation rates to be measured.

Many lines of evidence suggest that the main distinguishing feature of T Tauri stars is the copious flux of mechanical energy — winds, waves and shocks — which emanate from them. In some cases this is known to be a substantial fraction (10 per cent or more) of the radiative flux passing the photosphere. With this extraordinary result well established, it is perhaps surprising that the phenomenon has not

received more serious theoretical attention — what possible processes can lead to such a large proportion of the energy losses in mechanical forms? With Kuhi's important work perhaps hinting that rotation is not the primary cause of the effect, is it magnetic fields to which the theoreticians must now look for the explanation?

Participants stressed the desirability of coordinated observations at different wavelengths. In particular, the importance of space observations and the pressure of time on such facilities — like IUE and EXOSAT — made this vital. Possibilities for future coordinated observing programmes were discussed and it was concluded that May 1982 would provide the first major opportunity to deploy IUE and other facilities together and would be particularly suited for Southern stars. It was agreed that Kuhi of the Lick Observatory at Santa Cruz, California would be central coordinator and that any other interested parties would be welcome to join in coordinated campaigns. □

In the ensuing shakeout, three schools of biogeographical theory have emerged². The first is an updated version of Wallace's programme, centres of origin and mobile life, married with the mobile Earth of plate tectonic theory. The second, equilibrium theory, has grown from MacArthur and Wilson's 1967 book, *The Theory of Island Biogeography*. Equilibrium theorists apply the mathematical models of population genetics and ecology to biogeography, and investigate the chances of dispersal, colonisation and extinction empirically, through work on natural or artificial islands. The third school, vicariance biogeography, finds its roots in the voluminous and idiosyncratic prose of Leon Croizat³, one of whose slogans is 'Earth and life have evolved together'. According to vicariists, the distribution of life is governed not by the wanderings of animals and plants, but by Earth history. Tectonic events have split whole biotas, and historical relationships between areas may be reconstructed through the systematic relationships of species. Vicariance theory and method were reviewed at a major meeting in New York in 1979⁴.

A recent meeting* at the British Museum, set against this background, had spokesmen for all three schools. It did not develop into a showdown or conflict between the three; rather, speakers stuck close to their preferred school, with nods or sallies towards the others. Many speakers commented on the three approaches and the conflict between them, but no general argument or consensus emerged.

In view of the ideological fragmentation of the meeting, I mention only a few of the papers. Apologies are due to many speakers whose contributions are not mentioned, but some of them, botanists in particular, were still in the data-gathering rather than explanatory phase.

Amongst equilibrium theorists, E.C. Pielou (University of Dalhousie) contributed to the debate on punctuational change versus gradualism. She provided a method of analysis — ordered similarity matrices — which avoids qualitative judgments, and applied the method to forams from the Bay of Fundy. Step-like rather than gradual change was inferred in these forams, thought to be due to sudden change in tidal pattern. Similar results were reported by the palaeontologist R. Coope (Birmingham University), summarising his work on Quaternary and Pleistocene Palaearctic beetles. These animals show no detectable evolutionary change or extinction; instead the effect of climatic change is paramount. Coope believes his insect communities moved to dodge environmental change, rather than evolving or becoming extinct. D. Simberloff (Florida State University) presented a

Biogeography: in search of principles

from Colin Patterson

'BIOGEOGRAPHY is a strange discipline. In general, there are no institutes of biogeography; there are no departments of it. There are no professional biogeographers — no professors of it, no curators of it. It seems to have few traditions.' (see ref. 1). Biogeography in the accepted sense was invented just over a century ago, by Alfred Russel Wallace, who saw his *Geographical Distribution of Animals* (1876) as elaborating the two chapters on geographical distribution in *The Origin of Species*, just as Darwin had elaborated chapter 1 in *The Variation of Animals and Plants Under Domestication*. If biogeography has traditions, they are rooted in Wallace's works. As all traditions sooner or later must, these have come under attack during the past 15 years.

Wallace's programme was based in the geologist Dana's theory of fixed continents and permanent ocean basins, an idea endorsed by Darwin. If the continents and oceans have always been roughly where they are, and if 'every species has come into existence coincident both in space and time with a pre-existing closely-allied species' (Wallace's 1855 formulation of evolution),

the biogeographer's task is to infer when and where each species or group arose, and to explain the subsequent distribution of each. Help was sought in the fossil record, in anecdotal and historical records of dispersal, and in empirical investigation such as Darwin's immersion of seeds in salt water, and his scraping potential passengers from ducks' feet. The results of this programme became embodied in general theories: for example, competition and natural selection are more effective on large continental areas such as the northern landmasses (Darwin's 'more efficient workshops of the north'); rigorous climates promote evolution (the monoboreal hypothesis — life has radiated from the Arctic). According to these theories, the close relationship between the southern temperature biotas of Australasia, South America and Africa, evident to the botanist Joseph Hooker in the 1850s, was illusory; those regions were culs-de-sac where the oldest groups survived, forced southwards by the newer and more efficient products of northern workshops. These ideas, or variants of them, were dominant in biogeography until the middle 1960s, when geophysical evidence of continental mobility became overwhelming. With that came the realisation that the relationship between southern biotas reflected Earth history.

*As part of the British Museum (National History) centenary celebrations, the Museum joined with the Systematics Association in organising a week-long international symposium on biogeography, 6-10 April, under the title 'Time and Space in the Emergence of the Biosphere.' About 100 biologists and geologists met to hear almost 30 speakers.

Colin Patterson is in the Department of Palaeontology, British Museum (Natural History), London.

pessimistic summary of equilibrium theory, proposing that the equilibrium model is hard to define and difficult to test, but seems to have failed some tests. Simberloff was concerned with hypothesis testing, and the difficulty of using patterns to determine causes, themes which came up repeatedly.

I.R. Ball (University of Amsterdam) addressed the most general problems of biogeography and tried to steer a middle course between the shoals of various schools. Like Simberloff, Ball discussed hypothesis testing and the problem of distinguishing pattern from process, but he also spoke of recent emphasis (over-emphasis in his view) on Method. By Method, Ball meant the cladistic debate, which was mentioned as often at the meeting as in recent issues of *Nature*. The impact of cladistics on biogeography has been in two stages. The first, best exemplified in Brundin's 1966 monograph on austral chironomid midges⁵, used cladograms of taxa to estimate their centres of origin and routes of dispersal. In brief, the method is to correlate the cladogram and map and to use a parsimony procedure, minimising the number of dispersals, to localise branch points in the cladogram. The method is infallible, but is unduly sensitive to fossil localities and depends on the axiom that each taxon had a centre of origin from which it dispersed, by means which are assumed to depend in part on the capacities of the particular

organisms, and in part on chance. For example, by this method man is inferred to have originated in and dispersed from Africa, since our cladistic relatives, chimpanzees and gorillas, are found there, as are the most primitive man-like fossils.

The second stage of cladistic biogeography was developed within vicariance theory. This method seeks to combine cladograms of individual taxa into cladograms of areas⁶. In formal terms, cladistic analysis moves to a more general level, from that at which groups of organisms are the units, ordered by finding congruence between character cladograms (cladistic systematics), to one at which geographical areas of endemism are the units, ordered by finding congruence between cladograms of organisms inhabiting those areas. At this level of analysis, a cladogram of apes and humans would be of interest, not as a source of inferences about the areas of origin and routes of dispersal of those animals, but as one among many taxon cladograms bearing on the inter-relationships of the areas occupied by apes and humans. In this mode of analysis, centres of origin have the same status as ancestral species in cladistic systematics: they are regarded as irrelevant, because ancestors are components of trees, not of cladograms, and there are many trees compatible with one cladogram.

The problem with the application and testing of the method just described is shortage of the raw material, taxon cladograms. At the meeting, taxon cladograms were presented by several speakers, including J.R.G. Turner (Leeds University), Amazonian butterflies; J.D. Holloway (Commonwealth Institute of Entomology), Indo-Australian butterflies; and A.R. Milner (Birkbeck College, London), who found congruence between his cladogram of urodele amphibians and

inferred geological history of the northern continents. But the only attempt to combine cladograms of different taxa was by C.J. Humphries (BM(NH)), who took the southern beeches (*Nothofagus*) as his starting point, and arrived at a general cladogram of the southern continents. K. Thomsen (Aarhus University) applied the cladistic method and Rosa's theory of hologenesis to *Homo sapiens*, inferring a cosmopolitan ancestral population which was subdivided by successive vicariance events. This analysis engendered some heat amongst the anthropologists present. Our own species was also discussed by G. Nelson (American Museum of Natural History), who presented a historical survey of biogeographical thought from Linnaeus to the present, and used theories of the distribution of *Homo sapiens* to illustrate his theme that centre-of-origin biogeography is akin to Linnaeus's ideas on the location of Paradise.

My overall impression of the meeting was of a discipline in search of some discipline, or some principles. At least one speaker, Nelson, held out hope for the future, remarking that 'biogeography is still in its beginning phase, one comparable to astronomy at the time of Galileo'. To Nelson, the hope lies in area cladograms. To others, such as I.R. Ball, the hope is in a pluralistic approach. □

1. Nelson, G. J. *hist. Biol.* 11, 269 (1978).
2. Vuilleumier, F. C. r. somm. Séanc. Soc. Biogéogr. 54, 41 (1978).
3. Croizat, L. *Space, Time, Form: The Biological Synthesis* (Caracas, 1964).
4. Nelson, G. & Rosen, D.E. *Vicariance Biogeography: A Critique* (Columbia University Press, New York, 1981).
5. Brundin, L. K. *svenska Vetensk.Akad. Handl.* (4) 11, (1966).
6. Nelson, G. & Platnick, N.I. *Systematics and Biogeography: Cladistics and Vicariance* (Columbia University Press, New York, 1981).



100 years ago

THE GORILLA AND THE CHIMPANZEE

Mr. H. von Koppenfels, who is now engaged in explorations in the Gaboon district of Western Africa, in a letter published in the last number of the *American Naturalist*, states that he has good evidence of the existence of crosses between the male gorilla and the female chimpanzee. "This," he says, "settles all the questions about the gorilla, chimpanzee, Kooloo-Kmba, N'schigo, M'bouvé, the Sokos, Baboots, &c." Herr v. Koppenfels observes that the "French savants seem to have a special predilection for creating new species from variations in the form of the skull, such as often occur in this group of animals. The chimpanzee is found all over Tropical Africa, and naturally exhibits considerable variation. Du Chaillu's Kooloo-Kamba, N'schigo, and M'bouvé are not distinct species, and this traveller, who is certainly a man of merit, but is too credulous, has been imposed upon by the mendacity of the natives, which beggars description. The mongrel progeny of the male gorilla and female chimpanzee discovered by me is found but in individual cases, and as such deserves no special name."

from *Nature* 24, 30 June, 203, 1881.

Planet X: is it necessary?

from David W. Hughes

EPIHEMERIDES are tables recording the predicted positions of celestial objects against the sky background, and astronomers are still making strenuous efforts to improve their accuracy. Careful comparisons between, for example, the observed positions of Neptune and Uranus and the predicted positions of these planets were a prelude to the discovery of Pluto in 1930. It is thought by some astronomers that a continuation of this process, including Pluto, could now lead to the discovery of a 'Planet X', the next planet out. Others doubt the existence of more planets and believe that the tally of major planets is complete.

David W. Hughes is in the Department of Physics, University of Sheffield.

The last orbit of Pluto was published in 1972 and was obtained by analysing the positions of the planet recorded between 1914 and 1965. Over 500 observations were used, including 14 prediscovery ones where Pluto had been photographed but had not been recognized as a planet.

A new set of orbital parameters for Pluto has been computed by P.K. Seidelmann, G.H. Kaplan, K.F. Pulkkinen, E.J. Santoro and T.C. Van Flandern of the United States Naval Observatory, Washington DC (*Icarus* 44; 19, 1980). Another 15 years of observations have been added to the previous set, the majority coming from a Naval Observatory programme for observing Pluto, the same programme that fortuitously led to the discovery of

Pluto's satellite. This discovery and the analysis of the satellite's orbit removed a previously unknown quantity, the mass of Pluto, from the orbital calculation.

Pluto has now been observed over about a quarter of its orbit. The new results give the orbit a semi-major axis and period of 39.718 AU and 250.3 yr whereas Allen's *Astrophysical Quantities* (1973) quotes 39.44 AU and 247.7 yr respectively. The inclination is 17.139029° , the argument of perihelion 112.985694° and the eccentricity 0.25234645° , all for the epoch 1979 May 7.

Even now the residuals between the recent observations and the new ephemeris indicate that improvements must be made and to this end positional observations of the planet must continue.

Prediscovery estimates stated that Pluto's mass would be about five times that of the Earth. It is now known to be about a five-hundredth of the mass of the Earth. The ephemerides of Uranus and Neptune have been computed under the assumption that the mass of Pluto is very small. Comparison in right ascension and

declination have also been made between the observed positions of the two planets and the ephemerides. The results can be interpreted as indicating that a secular difference of the orders of half an arc second exists between observation and ephemeris for the orbits of both Uranus and Neptune.

The authors conclude that some unmodelled force is acting on the outer planets. Observations of six long-period comets, which have had more than one apparition, show that these too are being perturbed by an unmodelled force perpendicular to their orbital plane. The well known non-gravitational force acting on comets has no component in this direction. All these comets have aphelia beyond Neptune.

The cause of the force is unknown. That a Planet X beyond Pluto is exerting a gravitational effect is a possibility but the authors stress that the long and unsuccessful photographic search made by Clyde Tombaugh makes its existence unlikely. If, for example, Planet X was as

massive as Jupiter, Tombaugh concluded that he would have found it if it were closer than 470 AU. Also, this phantom 'Planet X' does not seem to have a measurable effect on the orbits of new comets. These comets, which have just been 'bumped' out of the Oort cloud by a passing star, seem to have near perfect parabolic orbits and there is no anisotropy in their major axis distribution. Furthermore, Planet X does not seem to affect the plane of symmetry of the Solar System. This seems to be dominated by Jupiter.

Our understanding of the motions of Uranus, Neptune and Pluto remain incomplete but the authors remind us that we should not leap to a 'Planet X' hypothesis too quickly. Leverrier made this error when, in 1859, he suggested that the anomalous advance of the perihelion of Mercury was due to perturbations exerted by a planet or planets inside Mercury. Even after the excitement caused by Lescarbault's supposed discovery of Vulcan, the correct explanation had to await Einstein's theory of relativity. □

Muscle regulation: a decade of the steric blocking model

from John Squire

THE first insights into the molecular mechanism of muscle regulation were obtained nearly ten years ago. The calcium concentration in the sarcoplasm, itself regulated by the nervous input to the muscle via the T-tubules and the sarcoplasmic reticulum, was known to be involved through the concentration-dependent binding of calcium to the actin filament protein troponin¹. However, little was known about how troponin influenced the structure of these filaments.

Evidence existed that troponin is associated with the rod-shaped α -helical protein tropomyosin¹, and that in the actin

filament the actin subunits form a structure which approximates to two coaxial long-pitched helical strands of globular beads² (Fig. 1). The tropomyosin molecules, each with a single troponin complex attached, were thought to run head-to-tail to form continuous protein strands along the two long-pitched grooves of the actin helix, such that each single tropomyosin/troponin repeat would be associated with seven actin subunits along each helical strand. But how could calcium binding to a single troponin molecule influence the binding of myosin heads to seven actin subunits? Clearly tropomyosin could be

involved, but in what way?

Part of the answer came when it was found that differences in the X-ray diffraction patterns of relaxed and contracting muscles³⁻⁶ could only be satisfactorily interpreted^{4,5,7} in terms of a change in position of tropomyosin in the grooves of the actin helix. Figures 1b, c and 2a illustrate the kind of movement involved. In relaxed muscle the tropomyosin (viewed end on in Fig. 2a) was thought to lie well out of the groove ($\theta = 45-55^\circ$), whereas in active muscle it moved much closer to the groove ($\theta = 65-75^\circ$). Later studies^{8,9} confirmed this tropomyosin movement, but how was it linked with thin filament regulation?

In 1970, Moore, Huxley and DeRosier¹⁰ obtained the first three-dimensional reconstructions from electron micrographs of negatively stained actin filaments decorated with the S-1 head sub-fragment of myosin. Their reconstructions could be interpreted to mean that the myosin heads attach to actin as shown in Fig. 2b. The observed tropomyosin shift was soon combined with this postulated myosin attachment position to produce the so-called steric blocking model of thin filament regulation^{4,5,7}. The 'off' position for tropomyosin in Fig. 2a is very close to the attachment site of myosin to actin in Fig. 2b, but the 'on' position is clear of this site. Thus tropomyosin could physically block the attachment of myosin to actin in

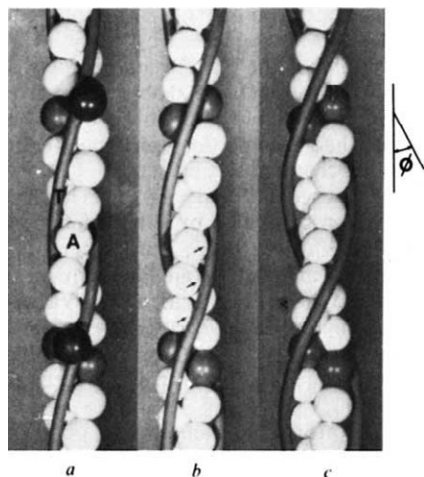


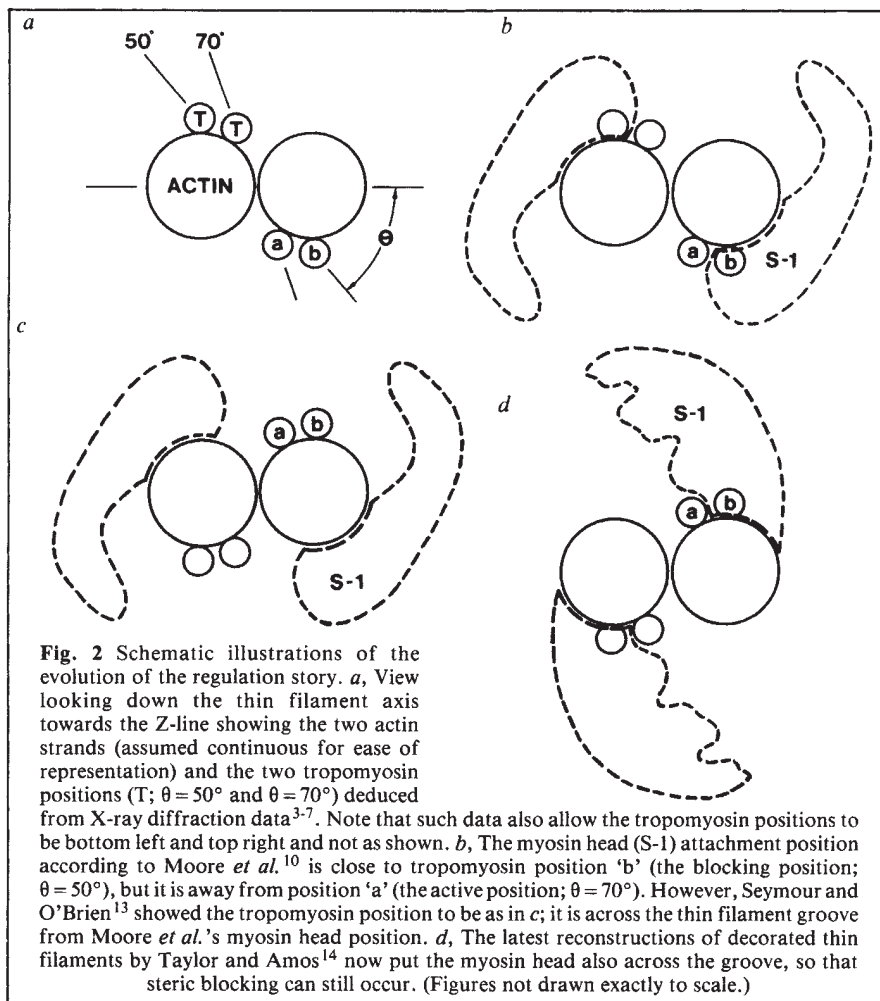
Fig. 1 *a*, Early model of the thin filament; actin monomers (A) form a helical structure² with tropomyosin strands (T) running centrally up the two long-pitch grooves of the helix¹. Troponin molecules (dark spheres) are attached to each tropomyosin molecule and there is one tropomyosin/troponin pair for every seven actin monomers. Grey actins mark the seven monomer repeat. *b* and *c*, Revised models^{4,5,7} in which the tropomyosin is not centrally placed but is at about $\theta = 70^\circ$ (*b*; see Fig. 2a), and about $\theta = 50^\circ$ (*c*). In the steric blocking model the tropomyosin position in *b* is to one side of the sites (arrows) where myosin heads can bind to actin; in *c*, these sites are blocked by tropomyosin. Troponin is not shown in *b* and *c*. According to Seymour and O'Brien¹³ the Z-line would be at the lower end of these filaments.

relaxed muscle (low calcium) but would roll into the groove and out of the way of the attachment site, due to calcium binding to troponin, when the muscle was stimulated. Although it was not known at the time that the myosin site and the tropomyosin strands were on the same side of the actin groove, this seemed a reasonable suggestion; otherwise such a steric blocking mechanism would be harder to envisage.

Nobody assumed that steric blocking would prove to be the whole story and biochemical evidence was soon obtained which suggested that additional aspects of thin filament regulation might be mediated through the actin subunits themselves, possibly as a result of tropomyosin movement^{11,12}. But it was not until 1980 that the steric blocking model received its first structural jolt. Seymour and O'Brien¹³, from studies of thin filament structure in I-segments, showed that tropomyosin does indeed lie on the opposite side of the actin groove to the myosin-binding site suggested by Moore *et al.*¹⁰ (Fig.2c). How could steric blocking occur now? Perhaps the biochemists were right and tropomyosin movement, about which there was little doubt, only acted indirectly through the actin subunits in modifying the binding site.

One of the inherent problems with three-dimensional reconstructions of helical structures such as actin filaments is that it is difficult to obtain reliable information about densities close to the axis of the helix. This is because, in the diffraction patterns used for the analysis, such information is at the outer edges (high radius positions along layer lines) where the effects of disorder in the specimen are most serious. For this reason Taylor and Amos¹⁴ have recently repeated the work of Moore's group¹⁰ using longer straight stretches of decorated thin filaments and more up-to-date methods, including minimal-dose electron microscopy. This enabled them to retain more information about the low-radius features in the decorated filaments — features relating to the actin subunits themselves — and they found that the peaks chosen to be actin by Moore *et al.*¹⁰ were probably part of the myosin head and that new density features close to the thin filament axis (Fig.2d) could more sensibly be assigned to actin.

These results clearly provide a new insight into the mode of attachment of the myosin head to actin. As early as 1965, Reedy, Holmes and Tregear¹⁵ suggested from electron micrograph appearances that in rigor muscle the long axis of the actin-bound myosin heads makes an angle (ϕ in Fig. 1) of about 45° to the filament axis. Combined with an apparent tilt closer to 90° in fixed relaxed muscle this result led eventually to the swinging crossbridge model of force generation in muscle¹⁶. However, modelling of X-ray diffraction



patterns of rigor muscles^{17,18} has generally tended to suggest an attachment angle (ϕ) rather closer to 90° than the 45° value suggested by Reedy *et al.*¹⁵. The shape of the myosin head itself makes the definition of tilt rather ambiguous; the head is probably curved, about 60 Å across proximal to the actin-binding site, tapering to < 30 Å at its distal end, and between 120 and 180 Å long^{14,19}. But the Taylor and Amos results suggest that ϕ is about 55–65° and a similar study²⁰ puts the angle closer to 75°. Although at first sight these values conflict with the appearances in Reedy's micrographs of rigor insect muscle^{15,21}, it should be remembered both that the S-2 link between the head and the myosin filament backbone is present in those sections and that until such micrographs are analysed in three dimensions only two-dimensional projections of the structure are being studied. In such views, the apparent tilt angle (ϕ) of myosin heads which do not have their long axes exactly parallel to the plane of sectioning will appear to be less (closer to 45°) than it is in three-dimensional reconstructions of decorated filaments where the angle can be measured more directly. Thus there may be no real conflict between these various results.

One of the most significant features of the recent Taylor and Amos work is that

the myosin attachment site (Fig.2d), like the tropomyosin position, has now also moved across the actin groove from the early model (Fig.2a). Tropomyosin and the myosin attachment site are close together once again and, although some aspects of regulation are clearly mediated by actin^{11,12}, the steric blocking model lives on. □

1. Ebashi, S. & Endo, M. *Prog. Biophys. molec. Biol.* **18**, 123 (1968).
2. Hanson, J. & Lowy, J. *J. molec. Biol.* **6**, 46 (1963).
3. Huxley, H. E. *Biochem. J.* **125**, 85 (1971).
4. Huxley, H. E. *Cold Spring Harb. Symp. quant. Biol.* **37**, 361 (1972).
5. Haselgrove, J. C. *Cold Spring Harb. Symp. quant. Biol.* **37**, 341 (1972).
6. Vibert, P. J., Haselgrove, J. C., Lowy, J. & Poulsen, F. R. *J. molec. Biol.* **71**, 757 (1972).
7. Parry, D. A. D. & Squire, J. M. *J. molec. Biol.* **75**, 33 (1973).
8. Spudich, J. A., Huxley, H. E. & Finch, J. T. *J. molec. Biol.* **72**, 619 (1972).
9. Wakabayashi, T., Huxley, H. E., Amos, L. A. & King, A. *J. molec. Biol.* **93**, 477 (1975).
10. Moore, P. B., Huxley, H. E. & DeRosier, D. J. *J. molec. Biol.* **50**, 279 (1970).
11. Bremel, R. D., Murray, J. M. & Weber, A. *Cold Spring Harb. Symp. quant. Biol.* **37**, 267 (1972).
12. Weber, A. & Murray, J. M. *Physiol. Rev.* **53**, 612 (1973).
13. Seymour, J. & O'Brien, E. J. *Nature* **283**, 680 (1980).
14. Taylor, K. A. & Amos, L. A. *J. molec. Biol.* **147**, 297 (1981).
15. Reedy, M. K., Holmes, K. C. & Tregear, R. T. *Nature* **207**, 1276 (1965).
16. Huxley, H. E. *Science* **164**, 1356 (1969).
17. Miller, A. & Tregear, R. T. *J. molec. Biol.* **70**, 85 (1972).
18. Holmes, K. C., Tregear, R. T. & Barrington-Leigh, J. *Proc. R. Soc. B207*, 13 (1980).
19. Elliott, A. & Offer, G. *J. molec. Biol.* **123**, 505 (1978).
20. Toyoshima, C. & Wakabayashi, T. *J. Biochem.* **86**, 1887 (1980).
21. Reedy, M. K. *J. molec. Biol.* **31**, 155 (1968).

John Squire is a Lecturer in Materials Science and leader of the Biopolymer Group at Imperial College, London.

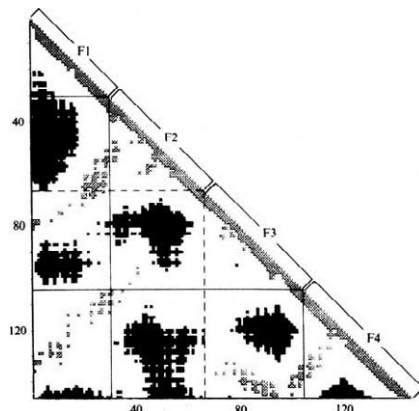
Exons and the structure, function and evolution of haemoglobin

from C.C.F. Blake

OVER the last few weeks there has been a flurry of papers in *Nature* that bear directly on the meaning of coding sequences or exons in the haemoglobin genes. In the mouse α - and β -globin genes there are three exons corresponding to amino acid residues 1–31, 32–99 and 100–141; and 1–30, 31–104 and 105–146, respectively. When the amino acid sequences of the α - and β -globins are aligned to maximize their structural homologies, their intron/exon junctions coincide. This immediately suggests that the 'meaning' of the exons is related in some way to the structure and function of the protein. The recent papers address themselves to the problem by analysing the exon-encoded regions in terms of the structure, function and evolution of the haemoglobin molecule.

Gö¹ has ingeniously used the diagonal plot (see the figure) to define regions of the polypeptide chain that are distant ($> 27\text{\AA}$) from one another in the globin fold. He has found that there are four such regions, which are not domains, but which can be called sub-domains, or perhaps 'compact structures'. When included on the diagonal plot the exon boundaries neatly divide off these compact structures from one another, with the exception of the large central exonic region which is composed of two compact structures. Because of this Gö suggested that the central exonic region might consist of two 'fused' exonic regions with a division somewhere between residues 66 and 71. Quite remarkably this prediction has been rapidly verified by Marcker and colleagues² (see this issue of *Nature*, p.677) in their determination of the structure of the leghaemoglobin gene from soybean. This gene is composed of four exons, corresponding to amino acid residues 1–32, 32–68, 69–103, 104–C terminus. Marcker's results suggest that exons are generally very stable, but that they can undergo fusion or separation. Gö's successful prediction of an exon on the basis of protein structure adds powerful additional support to the idea³ that they correspond to compact protein structures.

As Gilbert's suggestion⁴ of a relation between exons and protein functional units appears to have been validated in lysozyme⁵, consideration of this aspect of haemoglobin is of particular interest. In haemoglobin the functional correlation began with the suggestion⁶ that the central exonic region appeared to correspond to a haem-binding unit. This suggestion was



A plot of all the distances between α -carbon atoms in the haemoglobin β chain. Both ordinate and abscissa are the residue sequence number. Distances between pairs of atoms greater than $27\text{\AA}$ appear as dark regions. Solid lines are drawn at the joints between polypeptide segments (F1, F2 + F3, and F4 respectively) encoded by different exons — between residues 30 and 31, and 104 and 105. Note that these lines scarcely cross the dark regions.

analysed in detail by Eaton⁷, who not only confirmed this property of the central exon, but also showed that the distribution of intersubunit contacts showed a strong correlation with the exonic regions. Nearly all the $\alpha_1\beta_2$ contacts are in the central exonic region, and nearly all the $\alpha_1\beta_1$ contacts are in the third exonic region. This seems to suggest that while the contacts responsible for the early cooperativity of $\alpha_1\beta_2$ dimers may have developed within the haem-binding structure itself, the later tetramer cooperativity required the presence of the third coding sequence, which Eaton speculates may have replaced its predecessor by recombination. At the same time Beychok and his colleagues⁸ began an experimental examination of the properties of the central exonic region of human haemoglobin, which can be excised from the globin chain by clostripain treatment. Their studies first demonstrated that the separated central region was indeed capable of binding haem tightly and specifically. An extension⁹ of this analysis has now shown that although the central exon peptide can bind haem, it cannot stabilize the haem-dioxygen complex, which requires both the presence of the side exon products and the complementary subunit. The regainment of activity is accompanied by an increase in helical content, suggesting that a structural conformational change may be involved. With the verification of the haem-binding function of the central exonic region, it is interesting to consider the significance of

its division into two in leghaemoglobin. Despite Marcker's assertion that it has no obvious functional correlation, the division in fact neatly segregates the proximal and distal haem contacts. This suggests that the divided central exonic region may be a primitive form and that the two exons fused along the haemoglobin line to ensure a single protein product capable of specifically interacting with haem.

In a consideration of their results, Beychok and his colleagues note that the central exon product, isolated or in noncovalent association with the side peptide fragments, has spectral characteristics and other behaviour reminiscent of the *b*-type cytochromes. In 1975 Rossmann and Argos¹⁰ demonstrated a structural similarity between the globins and cytochrome *b*₅, which was later extended to include cytochrome *c*₅₅₁¹¹. In each case the closest agreement involved the haem-binding region and corresponds rather precisely to the central exon in the globin chain. When the protein folds were maximally superposed, the haem irons were found to be separated by $4.2\text{\AA}$ in the globin – cytochrome *b*₅ comparison and by $5.3\text{\AA}$ in the globin – cytochrome *c*₅₅₁ comparison, and the haem normals by 13° and 9° respectively. The results of these chemical and structural comparisons suggest that the globins and the cytochromes may have evolved from a common haem-binding domain encoded by one or more exons, which by combining with other gene elements has generated a multiplicity of haem-binding proteins of diverse function. These conclusions are very similar to those obtained from the examination of the exon structure–function relationships in hen and T4 lysozymes⁵, and lend strong support to Gilbert's proposal that in eukaryotes exons code for functional protein units which can serve in rapid protein evolution. The great puzzle now is why prokaryotic genes do not have this apparently advantageous structure. □

1. Gö, M. *Nature* **291**, 90 (1981).
2. Jensen, E.O., Paludan, K., Hyldig-Nielsen, J.J., Jørgensen, P. & Marcker, K.A. *Nature* **291**, 677 (1981).
3. Blake, C.C.F. *Nature* **273**, 267 (1978).
4. Gilbert, W. *Nature* **271**, 501 (1978).
5. Artymiuk, P.J., Blake, C.C.F. & Sippel, A.E. *Nature* **290**, 287 (1981).
6. Blake, C.C.F. *Nature* **277**, 598 (1979).
7. Eaton, W.A. *Nature* **284**, 183 (1980).
8. Craik, C.S., Buchman, S.R. & Beychok, S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1384 (1980).
9. Craik, C.S., Buchman, S.R. & Beychok, S. *Nature* **291**, 87 (1981).
10. Rossmann, M.G. & Argos, P. *J. biol. Chem.* **250**, 7525 (1975).
11. Argos, P. & Rossmann, M.G. *Biochemistry* **18**, 4951 (1979).

C.C.F. Blake is in the Laboratory of Molecular Biophysics, Department of Zoology, Oxford.

REVIEW ARTICLE

High-speed photography

George H. Lunn

Association for High-Speed Photography, 57 Whitedown Road, Basingstoke RG26 6BZ, UK

The most powerful of human senses is sight, to such an extent that to indicate we understand even an abstract concept we say "I can see that!". Until photography became a generally available technique, it required considerable skill and practice to produce any form of picture or illustration. The combination of variable time and pictures into one subject makes one appreciate the fascination of high-speed photography. Most high-speed photographers are users; they apply the techniques supplied to answer questions in some other study.

PHOTOGRAPHY began with Niepce and Fox-Talbot in the 1830s; indeed Fox-Talbot, by using an electric spark in his famous Royal Institute lecture (1851) simultaneously introduced high-speed photography. However, until 1879, when Ilford produced an emulsion coating machine, and 1880, when Kodak produced dry plates, photography was restricted to the few who could be bothered to work with wet plates. The first to capture public attention was Muybridge with his classical studies of animal motion, although Jansen in 1874 had studied the transit of Venus with a revolver camera. This was the period of cine-camera development (Edison, Lumiere, among others) with Marey and Bull in Paris creating mechanical cameras capable of 700 pictures per second (p.p.s.) (1894) and up to 2,000 p.p.s. (1904) with Worthington setting the pace in spark-photography studying splashes.

The early high-speed cameras used rotating disks carrying either lenses or film with the alternate one stationary to produce good short sequences such as the flight of a fly. Boys produced a rotating disk-bearing lens which much later was used by Schonland to show how a lightning strike behaved. These were limited in the number of pictures and top speed and restricted at that time by the low sensitivity of available photographic materials.

Rotating prism

It was not until 1932 that a suitable principle was developed to give an acceptable and commercial high-speed cine-camera. It was first used in the Olympic Games at Los Angeles as a photo-finish camera.

The standard cine-camera system moves film intermittently and only exposes it when it is standing still. This action severely limits the possible rates as neither the film nor the mechanism can suffer such repeated acceleration and deceleration beyond say 1,000 p.p.s. even with 16 mm film.

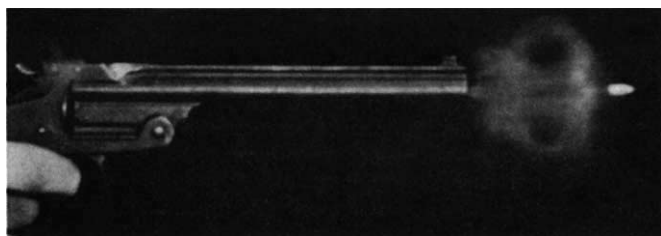
Allowing the film to move continuously and causing the image during exposure to move with the film produces 'image motion compensation', commonly called image compensation. This is achieved by placing a typically rectangular block of glass (prism) between the lens and the film and rotating it in synchrony with the film so that, over a useful angle, film and image move together. This method, first used successfully in 1932 in the Kirby race camera, was later taken over by Kodak, which produced the now famous Kodak high-speed camera using 16 mm film, 30 m long and taking 3,000 p.p.s. This was the standard camera in this field for almost 20 years. After 1945 it was overtaken by the Fastax camera which was lighter and produced in many models having a top speed of 8,000 full frame 16 mm p.p.s. with magazines up to 360 m. In recent years, other designs have taken over the market, usually with top speeds approaching 10,000 p.p.s. and 400-ft magazines but now with excellent speed control, synchronization and timing circuits using modern electronics.

This camera class has immense value in engineering and science, making clear many problems of the motion of machines and living subjects covering perhaps 95% of all high-speed applications.

This sequence shows a 0.22 bullet piercing a lightbulb. It starts in the top left-hand frame, with the bullet approaching the bulb and (moving clockwise) ends with the bullet having passed through and the bulb beginning to collapse. Photographs by Harold E. Edgerton of MIT, Cambridge, Massachusetts.

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Science Photo Library



Subject size

Deciding the camera speed required for any study involves some prior knowledge or estimation of the subject's velocity—not its actual velocity in metres per second but its relative velocity on film or the time it would take to traverse its image across the available picture. A suitable exposure time is one-thousandth of this time and the picture rate maximum would be the reciprocal of this exposure. This maximum is reduced by the shutter factor—the ratio of exposure time to frame interval.

Thus when recording a very large subject, the camera speed can be relatively low but with a microscopic object the camera speed is very high for the same subject velocity. Thus magnification in the camera magnifies the apparent velocity.

Synchronization

With 30 m of film at 10,000 p.p.s., the camera magazine is exhausted in half a second; thus the event and the camera must be arranged to coincide to at least 0.1 s or less: this is about the limit of human reaction. Thus synchronization must be automatic, involving a knowledge of both the camera and event variables. Skill in evolving methods of synchronization is the hallmark of a top high-speed operator, although he is much assisted by the electronics built into modern cameras.

Ballistics synchro

Ballistics synchro is a specialized application of these cameras with the compensating prism removed and a stationary slit close to the film plane. The camera is placed such that the event's motion is along the length of the film and actually imitates a focal plane shutter in which the film moves instead of the shutter. The result is identical to the current photo-finish cameras in which an isolated slit of space—the finishing line—is continuously photographed, thus making the one axis along the film equal to time, that is, a continuous study is made in time of the finishing line.

This is exemplified by a projectile photographed as it passes the camera. The result when the image and camera speeds are matched is an apparently normal picture of the projectile but as is seen in photo-finish pictures in which athletes or horses appear distorted, the change of time along one axis gives an unreal but useful image, especially of subjects whose synchronization is difficult.

Streak cameras

In principle streak cameras are also the same as photo-finish cameras but differ usually in being on shorter lengths of film—for example in early designs because the film was carried on a rotating drum and its was the length of the drum's circumference. Streak cameras are used for even higher speeds than the converted prism high-speed camera whose film speed maximum was 75 m s^{-1} . Although rotating drums are not very much faster than this, we will see later techniques that are.

Picture exposure time

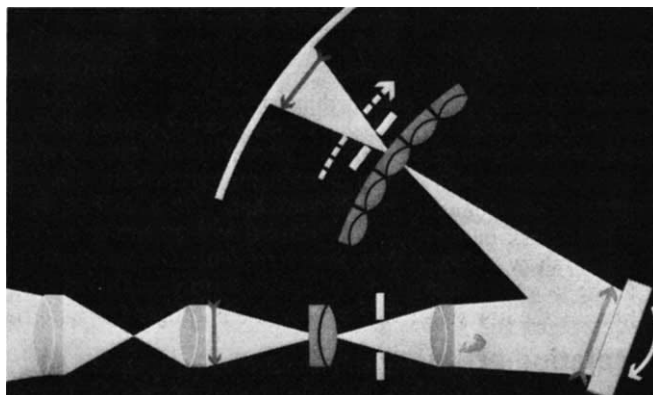
In the prism camera in its normal mode with image-compensation the maximum single picture exposure time is obviously something less than the frame interval whose minimum is $1 : 10,000$ of a second or $100 \mu\text{s}$. By inserting a slit near the film creating a stationary focal plane shutter, the individual exposure can be as short as $1 \mu\text{s}$. In any normal single picture camera, mechanical shutters cannot readily go below 1 ms. Thus for individual pictures with very short exposures, other shutters are necessary.

There have been various 'explosive' shutters which are only capable of being used once but they are really a last resort and are only used in very specialized and unique experiments. Over the same century that high-speed cameras developed, the Kerr-cell shutter has evolved from Kerr's discovery of the electro-optic effect in 1875.

Kerr discovered that some liquids in an electric field displayed birefringence, that is, it had two different refractive indices to light vibrating in orthogonal planes, and would thus have different transit times through the same pathlength. Plane-polarized light in the correct conditions would appear to have rotated its plane of polarization while traversing the cell. (The plane polarized light will have been resolved into two beams at $\pm 45^\circ$; one portion would have been retarded $\lambda/2$, the new combination appearing as having rotated 90° .)

We have now exchanged the motion of solid bodies—shutters, mirrors, prisms, drums—for the motion of molecules in a liquid. The final limitation occurred with solid bodies when they broke. In this case the limit is set by the molecular relaxation time typically of the order of 10^{-10} s. This cannot be fully achieved but we have reached a shutter action of 1 ns, with the flexibility of electronics to achieve any longer time. The Faraday magneto-electric effect is also usable but is electrically limited as it is necessary rapidly to change an inductive circuit as opposed to a capacitive circuit in the Kerr effect.

Both suffer from light losses as it is necessary to use polarizers giving a maximum transmission of 10% with a cutoff down by a factor of 10^{-4} . They have both been applied successfully and can give high quality images.



The rotating mirror framing camera. The film stands still while the rotating mirror moves. Image motion is compensated optically, and maximum speed (8 mm image) is about 10 million pictures per second.

Electronic flash

The first short exposure pictures (Fox-Talbot, Worthington) were by the light from a spark discharge. This was refined around 1932 by Edgerton, who put the spark in a bottle in an inert gas such as xenon and produced a brighter, more reliable light source from the same energy input. The flash duration is restricted to not less than $1 \mu\text{s}$ (in special high voltage equipment) and is in low voltage, amateur equipment about 1 ms, although the new computer flash units can be as short as $30 \mu\text{s}$. Edgerton extended this to stroboscopes, a series of short duration flashes, with readily available units giving 2,500 flashes per second, each flash being about $1 \mu\text{s}$, with longer, brighter flashes at lower repetition rates.

Most people who have a camera have also used an electronic flash, usually without fully appreciating its short exposure capability. To see Edgerton's pictures is really to understand what flash can do.

Sparks have been made with durations as light sources down to a few nanoseconds, the advantage being that the ionized gas which created the light can deionize and dissipate its heat sufficiently quickly whereas a discharge in gases like xenon in a bottle deionizes and dissipates more slowly.

Rotating mirror cameras

To achieve higher camera rates by mechanical means required a new means of image-compensation. Rotating mirrors to move a beam of light along a film arc had been used many times before. A good example of this was given by Blondlot (1888) who used observation through a rotating mirror to attempt to measure the time taken for the Kerr effect, his estimated minimum being 2.5×10^{-6} s, while Abraham and Lemoine (1899) measured an optical time delay of 80 cm which equals 2.7×10^{-9} s.

Miller (1939) proposed a rotating mirror framing camera system which was not really used until 1944 and even then was not totally appreciated. The basic idea is simple. You have a real image at or near the surface of a thin flat mirror which, rotating, sends the image beam to an arc of small lenses, each of which produces an image on a film arc. The light beam moves but the image detail on the film does not move because it is an image of a real image very close to the centre of rotation.

This system in a British camera (AWRE design, Barr and Stroud manufacture, JHPI distribution) can take pictures each of 8 mm diameter at rates up to 8×10^6 per s. At that rate, the mirror spins at about 5 kHz, 80% of the speed at which its own centrifugal forces would pull it apart, its outer edge moving above Mach 1. This top speed can only be marginally improved using titanium and as it was shown by Schardin (1956) that diffraction limits the practical optical beam length, these figures are as close to physical limits as it is reasonable to go.

Such a rotating mirror can equally well move a slit image along the film producing the classical streak camera with a writing speed at the film of $40 \text{ mm } \mu\text{s}^{-1}$ and a reasonable time resolution of 2.5 ns, which shows the excellence of nineteenth century workers. Framing and streak cameras have proved invaluable in studying explosives, properties of materials, aerodynamics, atomic fission and fusion, plasma physics and many other fields.

Having exhausted or at least reached the physical limits of solids, liquids and molecules, the only remaining hope for higher speeds was atomic particles: electrons.

Photoelectric devices

The photoelectric effect was discovered by Hertz (1887), explained by Einstein (1905) and for our purpose put to good effect by Holst and colleagues in 1934 when the first image tube was made. An image tube in its simplest form is a plane photocathode onto which an image is focused and a phosphor screen parallel and in close proximity to the photocathode with a suitable electrical potential between them such that electrons pass in direct paths one to the other and cause a visible image to appear on the screen. This is the first form to be attempted but it was not an immediate success because vacuum and clean room techniques had not been developed sufficiently.

The first good tube was produced at Mullards (1950) with a long electron path and used electromagnetic focusing. By puls-

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Another Edgerton photograph, this showing a bullet fired through a row of balloons with obvious results.

ing its tube voltage (6 kV) on and off, it operated as a shutter down to 10^{-7} s and as a streak camera (a slit image at the photocathode) with electromagnetic scanning, it gave an image on film at $100 \text{ mm } \mu\text{s}^{-1}$ and a possible time resolution below 1 ns. Improved technology now produces a wide variety of image tubes with proximity diodes and electrostatically scanned and focused tubes.

The leading camera system commercially is the Imacon (JHPI) which can be a framing camera up to 600×10^6 per s (5–15 frames) or a streak camera with resolution times of a few picoseconds. In a ballistics-synchro mode, its optical gain is enough to permit photography of a bullet in flight using only sunlight. These cameras can be made with response to IR, visual, UV or X-ray inputs. The picture output can be through a lens and camera or by contact using a fibre-optic face plate at the screen. Fibre-optic input or coupling in some image tubes is also possible.

Image intensification is possible with normal intensifiers, in cascade if desired, or even more conveniently by channel plate intensifier. This is a relatively new device (Adams and Manley, 1965) which has combined fibre-optic techniques with secondary emitters in an original manner and by a complex and clever technology with a British-French-Dutch combination through Philips. Take a fibre-optic face plate which is a mosaic of short glass rods coated by another different glass. These rods will be $40 \mu\text{m}$ diameter and 2.4 mm long, making a face plate 2.4 mm thick and 25–50 mm diameter. The glass of the rods has been selected so that it can be dissolved away leaving a face plate of tubes, the coating glass not being soluble. The inside of every tube is now coated with a secondary emitter and the end faces plated to give electrical connection. The plate is encased in an evacuated tube with a potential from face to face and a photocathode in proximity to one side and a phosphor at the other. Electrons emitted by light on the photocathode will enter the channel plate tubes and suffer numerous collisions with the secondary emitter surface, producing a shower of electrons to pass to the phosphor. Electron gains up to $\times 10^5$ are possible equivalent to the same degree of image intensification and making either very faint subjects visible or giving enough signal from a very short exposure to produce a recordable image.

Coupling these devices together in various systems produces instruments which are barely credible and were certainly not imaginable a few years ago. Coupling these devices with the current advances in lasers whose very short pulses are only measurable by these newest image tubes and it is possible that we may see the dream of safe energy from fusion come true.

As almost none of these devices was reasonably imaginable at the beginning of high-speed photography, it is foolish almost to speculate. One thing is certain, we have repeatedly arrived at a physical limitation of a system. When we remember that the frequency of light is only 10^{15} Hz and lasers and image tubes are already at 10^{-12} s, there is little room for further advance except perhaps by changing to higher radiation frequencies than light. The image tube chapter is only 30 yr old, and yet time resolution has moved 4 orders (10^{-8} – 10^{-12}). Could another 30 years show that even time is not divisible continuously?

Recent developments

The micro-channel plate multipliers produced under Philips with French and British cooperation have made a considerable impact in the image-tube field as they can be coupled with other tubes or incorporated as part of another tube bringing high gain intensification in an adaptable form.

In a simpler field, at camera rates above normal, are two 'convenience' systems. First the Polavision Super-8 cine-system by Polaroid-Land gives a camera capable of speeds up to 300 p.p.s. with immediate on-the-spot processing and playback 2 min after taking. It can be monochrome or colour, very portable, flexible and simpler to use—ideal for machine fault finding in automated factories.

Second, Video is now capable of leaving the real-time range and giving high speed taking with normal, still frame or low speed playback. Although not yet on the market the Spin Physics SP2000 system claims full frame rates up to 2,000 p.p.s., playback, duplication ability and x-y analysis.

ARTICLES

Gravitational focusing and the association of distant quasars with foreground galaxies

Claude R. Canizares

Department of Physics and Center for Space Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

The surface density of quasars brighter than a given threshold is enhanced near a galaxy by the gravitational focusing of light from those quasars that are nearly co-aligned with individual stars in the galaxy halo. The focusing can intensify and make detectable quasars which would otherwise have been below the threshold. If galaxies have massive haloes of macroscopic objects then the resulting enhancement could explain the apparent association between quasars and galaxies with discrepant redshifts found in a statistical analysis. Therefore, such quasar-galaxy associations need not require the abandonment of the cosmological interpretation of the redshift.

THERE have been numerous discoveries of quasars with moderate to high redshifts located near galaxies with very small redshifts (see refs 1, 2 and references therein). Although the statistical interpretation of each of these discoveries is fraught with systematic uncertainties and is therefore controversial, the number of proximate quasar-galaxy pairs may indeed exceed that expected from the superposition of two uncorrelated populations. More global tests of entire samples give small, marginally significant indications of a positive correlation between the positions of quasars and galaxies^{3,4}. The controversy surrounding all these findings is exacerbated by the implication that a positional association between galaxies and quasars necessarily implies that at least some of the quasars are at the distances of their neighbour galaxies and thus that their redshifts do not reflect the Hubble flow but require a noncosmological explanation².

I suggest here that the apparent surface density of quasars brighter than some threshold is in fact enhanced near a galaxy. This enhancement is due to the gravitational focusing of the light of quasars that are nearly coaligned by chance with individual stars in the galaxy halo. The focusing intensifies and makes detectable quasars whose brightness would otherwise fall below threshold. Plausible estimates of the relevant parameters imply density enhancements which are sufficient to account for the positive correlation between quasar and galaxy samples.

Gott and Gunn⁵ have previously suggested that the quasar surface density enhancement due to gravitational focusing by an entire galaxy could explain a close pair of quasars; Barnothy and Barnothy^{6,7} propose that all quasars are gravitationally intensified and that focusing by globular clusters enhances the quasar density near galaxies; Chang and Refsdal⁸ consider the gravitational effect of a star on the quasar image formed by a galactic gravitational lens. The latter topic is also treated by Gott⁹ in a paper received after my initial submission of this article. His conclusions regarding the dependence of the variability time scale on the mass of the focusing star are similar to those presented here. Finally, Turner¹⁰ has recently considered the effects of gravitational intensification by galaxies on various measures of quasar evolution.

Gravitational focusing effects by stars on quasar surface density

The enhancement in the quasar surface density due to an assemblage of stars (point masses) is calculated in two steps. First I find the probability that a given line-of-sight through the assemblage passes close enough to a star so that its action as a gravitational lens will intensify a background light source by an amount $>\mu$. Then I fold this with the surface density of background quasars that are below the detection threshold by a

factor μ and integrate over μ to derive the total enhanced density.

Einstein¹¹ and Refsdal^{12,13} give expressions for the intensification μ due to the gravitational focusing of light by an object B of mass M_* located near the line of sight from the observer O to the source S:

$$\mu(x) = \frac{1}{2} \left[\frac{1}{x} (1+x^2)^{1/2} + x(1+x^2)^{-1/2} \right] \quad (1)$$

where $\mu \equiv f/f_0$ and $x \equiv \beta/\alpha_0$. Here f is the flux seen by the observer, f_0 is the flux he would see in the absence of focusing, β is the angle between the true positions of B and S measured from O. Also

$$\alpha_0 \approx 4 \left(\frac{GM_*}{c^2} \right)^{1/2} A_B^{-1/2} = \alpha_{11} \left(\frac{M_*}{A_B} \right)^{1/2}$$

with

$$\begin{aligned} \alpha_{11} &= 8.9 \times 10^{-10} \text{ rad } M_\odot^{-1/2} \text{ Mpc}^{1/2} \\ &= 1.8 \times 10^{-4} \text{ s } M_\odot^{-1/2} \text{ Mpc}^{1/2} \end{aligned}$$

Here A_B is the distance from O to B, whose redshift is assumed small. I also assume that A_B is much less than the distance from O to S and that the splitting of the image is unresolved.

From equation (1) μ is >1.06 for $x < 1.0$ or $\beta < \alpha_0$. Thus the light of a quasar will be noticeably intensified if it falls within a solid angle $\sim \pi \alpha_0^2 = 1.0 \times 10^{-7} (M_*/M_\odot) (A_B/1 \text{ Mpc})^{-1} \text{ s}^2$ of a star of mass M_* at distance A_B .

The effect of gravitational focusing on the apparent surface density of quasars near galaxies is calculated as follows. The differential surface density of stars with mass M_* in a galaxy is given by the expression $(dN_*/d\Omega)\psi(M_*) dM_*$ where $dN_*/d\Omega$ is the surface density of all stars at some point in the galaxy and $\psi(M_*)$ is the mass distribution function, which I assume is the same throughout the galaxy. Then the probability density that a line of sight from Earth through the galaxy passes within x to $x+dx$ of a star with mass between M_* and M_*+dM_* is

$$p(x, M_*) dx dM_* = 2\pi \alpha_{11}^2 A_B^{-1} \left(\frac{dN_*}{d\Omega} \right) M_* \psi(M_*) dM_* x dx$$

Integrating over all M_* gives

$$p(x) dx = 2\pi \alpha_{11}^2 A_B^{-1} \frac{dN_*}{d\Omega} \langle M_* \rangle x dx \quad (2)$$

where $\langle M_* \rangle$ is the mean stellar mass in the galaxy. Equations (1) and (2) can be used to transform $p(x)$ to $p(\mu)$, which can then be integrated to derive the probability $P(\mu)$ that a given line of sight will result in an intensification $>\mu$:

$$P(\mu) = \int_{\mu}^{\infty} p(\mu) d\mu = \frac{\pi}{2} \frac{\alpha_{11}^2}{A_B} \Sigma_M [\mu(\mu^2 - 1)^{-1/2} - 1]$$

where $\Sigma_M = \langle M_* \rangle dN_*/d\Omega$ is the surface density of mass (per solid angle) in the galaxy measured along the line of sight.

$\Phi_\theta(f_i)$ is the angular surface density of quasars with apparent flux greater than a threshold f_i at some angular distance θ from the centre of a galaxy that has an outer radius R and is at a distance A_B . The enhanced density is related to the surface density of quasars brighter than f_i in the absence of focusing, $\Phi_0(f_i)$, by the expression

$$\Phi_\theta(f_i) = \Phi_0(f_i) + \int_0^{f_i} \Phi'_\theta(f) P\left(\frac{f_i}{f}\right) df \quad (3)$$

for $\theta < R/A_B$. Here $\Phi'_\theta(f)$ is the derivative of $\Phi_0(f)$. The second term gives the enhancement due to the gravitational intensification of quasars which would otherwise fall below the detection threshold.

I take a quasar surface density of the form $\Phi_0(f) = \phi_0(f/f')^{-\eta}$ for $f > f_{\min}$ where η , ϕ_0 and f' are constants. Below I will piece together a surface density distribution from several segments of this form. The integration of equation (3) then yields

$$\Phi_\theta(f_i) = \Phi_0(f_i) [1 + \zeta(\theta)] \quad (4)$$

where

$$\zeta(\theta) = F(\eta; \mu_{\max}) (\pi/2) \alpha_{11}^2 A_B^{-1} \Sigma_M(\theta) \quad (5)$$

Here $F(\eta; \mu_{\max})$ is an expression which generally depends on both η and $\mu_{\max} = f_i/f_{\min}$, the intensification required to bring the faintest quasars above threshold. However, the $\mu_{\max}$ dependence is important only when $\eta > 2$, because then the rise in the surface density of faint quasars compensates the fall in the probability of large intensification, so the faintest quasars dominate. Where $F(1; \mu_{\max}) = 1$ and $F(1.5, \mu_{\max}) = 2.3$, one finds $F(2; \mu_{\max}) = 1/2 + \ln(2\mu_{\max})$, and $F(2.5; \mu_{\max}) = (5/2)(\mu_{\max}^{1/2} - 0.5)$ for $\mu_{\max} \geq 1.5$.

Equation (5) can be recast in simple terms if the mass distribution of the galactic halo approximates a self-gravitating isothermal sphere out to some radius R . Then away from the core but inside R one has $\Sigma_M(\theta) = A_B \sigma_v^2 / 2G\theta$, where σ_v is the line-of-sight velocity dispersion in the halo⁹. This gives

$$\begin{aligned} \zeta(\theta) &= \frac{\pi F \alpha_{11}^2 \sigma_v^2}{4G\theta} \\ &= 4\pi F \left(\frac{\sigma_v}{c}\right)^2 \frac{1}{\theta}, \quad \text{for } \theta < R/A \end{aligned} \quad (6)$$

Finally, with equation (6), equation (4) can be integrated over a solid angle $\Omega = \pi\theta_{\max}^2$ centred on the galaxy to find $\langle n \rangle_\Omega$, the expected number of quasars that are above the threshold within the solid angle, in terms of $\langle n \rangle_0$, the expected number in the absence of gravitational focusing.

$$\begin{aligned} \langle n \rangle_\Omega &\approx \langle n \rangle_0 \left[1 + \frac{1}{\Omega} \int \zeta(\theta) d\Omega \right] \\ &\approx \langle n \rangle_0 [1 + 2\zeta(\theta_{\max})], \quad \text{for } (\theta_{\max} < R/A_B) \end{aligned} \quad (7)$$

with $\zeta(\theta)$ given by equation (6).

To estimate parameters I first consider $F(\eta; \mu_{\max})$. The value of η is found by comparing measurements of quasar surface densities at different limiting magnitudes. These measurements are generally uncertain by up to a factor of 10 (refs 15, 16) but there is some indication that η varies from $\eta \sim 2.5$ at limiting magnitudes ~ 16 – 18 mag (ref. 17), through $\eta \sim 2.0$ at 19 mag (ref. 18) to $\eta \sim 1.5$ at ~ 21 – 22 mag (refs 15, 16, 18, 19 and M. Schmidt, personal communication). Taking a piecewise continuous quasar surface density with these values of η and breaks at 18 and 21 mag gives the dependence of F on limiting magnitude shown in Fig. 1.

The line-of-sight velocity dispersion, σ_v , in the outer regions of a galaxy halo can be estimated from measured velocity dispersions and from rotation curves using assumed dis-

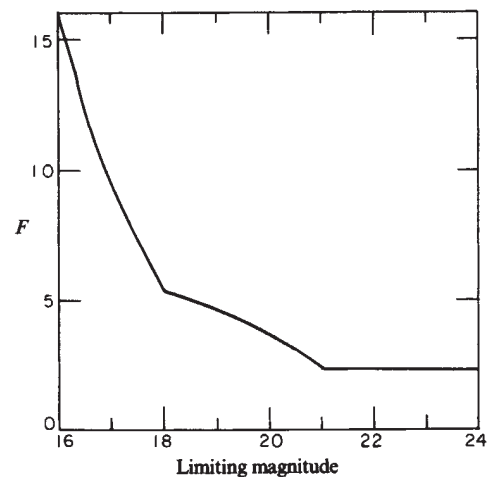


Fig. 1 Approximate value of the factor F , which appears in expressions for the enhanced quasar surface density, compared with the limiting quasar magnitude. Here F has been computed for a piecewise continuous unenhanced quasar surface density of the power-law form given in the text with exponent $\eta = 2.5$ for limiting magnitudes $m < 18$, $\eta = 2.0$ for $18 < m < 21$, and $\eta = 1.5$ for $m > 21$.

tributions of the galactic mass^{20–22}. In the nuclear regions of elliptical galaxies one observes dispersions of 200 – 400 km s^{-1} (ref. 23); the highest values are for the most luminous galaxies. In the absence of additional information one assumes that the halo and nuclear velocity dispersions are similar, although Dressler²⁴ found that for a cD galaxy the dispersion rises to $\sim 500 \text{ km s}^{-1}$ in the envelope. For the spiral galaxies, measurements of bulge velocity dispersions and of rotation curves yield estimates of $\sigma_v < 250 \text{ km s}^{-1}$ for the haloes^{20,22}; again σ_v increases with galaxy luminosity.

Scaling equation (6) to representative values of the parameters gives

$$\zeta(\theta) = 0.22 \left(\frac{F}{5}\right) \left(\frac{\sigma_v}{300 \text{ km s}^{-1}}\right)^2 \left(\frac{\theta}{\text{min}}\right)^{-1} \quad (8)$$

Thus if galactic haloes are made of macroscopic objects, the surface density of quasars seen through the haloes can be significantly enhanced. For bright quasars and massive galaxies one can have values of ζ four or five times larger than the representative value in equation (8), giving from equation (7) a doubling of the expected number of quasars within 2 arc min of the galaxy. Of course, our ignorance about the true quantity and distribution of mass in galactic haloes and about the surface densities of faint quasars means that the numerical values of ζ should be treated as approximate estimates of the true enhancement factor.

The above analysis neglects the gravitational focusing due to the distributed, total mass of the galaxy. One can show from the equations of refs 5, 9–14 that the enhancement in quasar surface density for the galaxy, ζ_g , depends on the 'critical angle' $\theta_c \approx 8\sigma_v^2/c^2$. For $\sigma_v = 300 \text{ km s}^{-1}$, $\theta_c = 1.7''$. At angles θ such that $R/A_B > \theta \gg \theta_c$ one has $\zeta_g(\theta) \approx (\eta - 1)\theta_c/\theta$, where η is the exponent in the quasar surface density distribution at the limiting magnitude of interest (because the intensification by the galaxy as a whole is small for $\theta \gg \theta_c$, the enhancement is due to quasars only slightly fainter than the threshold). If $\eta = 1$, the rise in the surface density of faint quasars just compensates the loss in solid angle caused by the magnification of the galaxian gravitational lens, and $\zeta_g = 0$. The ratio of $\zeta_g(\theta)$ to $\zeta(\theta)$ for stellar focusing from equation (6) is

$$\frac{\zeta_g}{\zeta} = \frac{2(\eta - 1)}{\pi F}$$

This is never greater than ~ 0.15 for the parameters specified above, so ζ_g can be safely neglected in the present work. When

$\theta \sim \theta_c$, the galaxian focusing can dominate; this is the relevant regime for discussion of the double quasar¹⁴ and is considered by Chang and Refsdal⁸ and Gott⁹. Note that a globular cluster has $\sigma_v < 10 \text{ km s}^{-1}$ (ref. 25) and $\theta_c < 0.002''$ so it has a negligible effect on quasar surface densities, contrary to the claim of ref. 7.

It is also necessary to consider the importance of the finite size of the quasar emission region (the finite size of the intervening star has no effect for $M_* \gg (10^6 \rho^{-2} A_B^{-3}) g$, where ρ is its density in g cm^{-3} and A_B is in Mpc). The size scale at the quasar is given by $l_0 \sim A_s \alpha_0$. Taking A_s , the 'angular diameter distance' to S, from the expressions of Young *et al.*¹⁴ gives

$$l_0 \sim 0.4 \left(\frac{H_0}{100} \right) \left(\frac{M_*}{M_\odot} \right)^{1/2} \left(\frac{A_B}{10 \text{ Mpc}} \right)^{-1/2} \text{ pc}$$

for quasar redshifts > 0.5 . Here H_0 is Hubble's constant in $\text{km s}^{-1} \text{ Mpc}^{-1}$.

The dimensions of the continuum emitting regions of quasars are thought to be $\ll l_0$ (ref. 26), so equations (5) and (6) are directly applicable to them. The regions of broad-line emission probably have dimensions $\sim l_0$ unless $M_* \ll 1 M_\odot$. One can show that an extended source of uniform surface brightness and dimension l will undergo an intensification $\mu_e \sim [1 + (l_0/l)^2]^{1/2}$ if it is co-aligned with the star. As it is possible to have $\mu > \mu_e$, the broad emission lines of quasars with large continuum region intensifications could appear anomalously weak. The dimensions of the narrow line emitting regions of quasars are almost certainly $> l_0$ (ref. 26), and so these lines would undergo negligible intensification as would radio or IR emission from extended regions.

The mass M_* of the individual objects serving as gravitational lenses enters into the expression for l_0 and for the variability time scale discussed below but does not appear explicitly in the quasar surface density enhancement of equation (5). The latter equation depends only on the mass surface density in the intervening galaxy. (In fact, one can use equation (5) to recover the order of magnitude of Press and Gunn's²⁷ result that the enhancement in quasar surface density for a universe with a uniform distribution of point masses is approximately equal to Ω_0 , the ratio of the density of the point masses to the cosmologically critical density; see also ref. 9). The validity of the single scattering approximation on which equation (5) is based is also independent of M_* (the probability that a line of sight falls within $\pi \alpha_0^2$ of a star in a galaxy halo is $8\pi \sigma_v^2 / c^2 \theta = 0.09$ for $\sigma_v = 300 \text{ km s}^{-1}$ and $\theta = 1'$). However, equation (1) is based on geometrical optics^{11,28} which requires that the individual objects have dimensions that are large compared with the wavelength of the focused radiation. If galactic haloes are composed largely of massive neutrinos, then they would not contribute to equation (5) and the value of ζ would be substantially reduced.

Comparison with observations

I now compare the estimates of the surface density enhancement of quasars near galaxies with the observational evidence of Seldner and Peebles³ and Burbidge¹. This is done by deducing from the observational data those values of ζ which would account for the apparent quasar-galaxy correlations.

Seldner and Peebles³ find a statistically significant enhancement in the galaxy-quasar two-point angular correlation function. The enhancement is equivalent to that which would occur if there were 570 ± 150 extra galaxies added to the $\sim 5,300$ galaxies expected by chance within $17.5'$ of all the quasars in the sample. Their result has the virtue of being completely independent of the assumed value of the mean quasar surface density, although it is certainly affected by selection effects as they discuss.

Because of the galaxy-galaxy two-point correlations²⁹, a real association between a quasar and a single galaxy will give rise to an apparent correlation between the quasar and other galaxies as well. Within some distance δ of the quasar (and its associated galaxy) one would find an excess number of galaxies given by

$$N_e \approx 1 + n_g \int_0^\delta (w_{gg}(\delta) - 1) d\Omega(\delta)$$

where n_g is the mean galaxy density³⁰ and w_{gg} is the galaxy-galaxy two-point correlation function²⁹. Evaluating this for $\delta = 17.5'$ gives $N_e = 75$ so that only ~ 8 of the 382 quasars in the Seldner and Peebles sample need have a real association with a galaxy to account for the excess quasar-galaxy correlation. This is consistent with Seldner and Peebles' finding that the enhanced quasar-galaxy correlation disappears when the 10 largest contributors to the enhancement are removed. In fact, 3 of those 10 are on Burbidge's list of quasar-galaxy pairs, which further suggests that the excess correlation could be due to a small number of 'real' associations.

It is possible to make an estimate of $\bar{\zeta}_{sp}$, the implied value of the average quasar density enhancement factor for the galaxies in Seldner and Peebles' sample that is needed to account for the ~ 8 true quasar-galaxy associations. The total number of observed quasars N_Q is given by

$$N_Q \equiv N(1 + 2\bar{\zeta}_{sp}\Omega_G/\Omega_T) \quad (9)$$

where N is the number in the absence of focusing and Ω_G/Ω_T is the fractional solid angle subtended by galaxies. The typical galaxy in the Shane-Wirtanen sample used by Seldner and Peebles has redshift ~ 0.08 (ref. 3) so $A_B \sim 240 \text{ Mpc}$ for $H_0 = 100 \text{ km s}^{-1} \text{ Mpc}^{-1}$. The galaxy luminosity function in a sample limited by apparent magnitude is peaked near $L^* = 2.5 \times 10^{10} L_\odot$ (ref. 31) and so has $\sigma_v \sim 300 \text{ km s}^{-1}$ (ref. 23). The outer radius R of the halo of such a galaxy is probably ~ 3 times its Holmberg radius of 40 kpc (refs 21, 32). Thus a typical galaxy subtends $2.4 \times 10^{-3} \text{ deg}^2$, and the mean galaxy density is $\sim 50 \text{ deg}^{-2}$ (ref. 30), so $\Omega_G/\Omega_T \sim 0.1$. Also $N_Q = 380$, $N \leq N_Q$, and I require $2N\bar{\zeta}_{sp}\Omega_G/\Omega_T \sim 8$. Inverting this last expression gives $\bar{\zeta}_{sp} \sim 0.1$. This compares remarkably well to equation (8), which with nominal values for F , and σ_v , and with $\theta = R/A_B \sim 1.7'$ gives $\zeta = 0.13$.

Burbidge¹ has made a statistical analysis of specific galaxy-quasar associations found by Arp and others. For separations $\leq 3'$ there is an excess of observed over predicted quasars nearly all of which is due to the excess between $1'$ and $2'$. For quasar-galaxy separations of $3'$ – $10'$ there is a deficit of observed quasars; presumably selection effects have caused many to be overlooked. The predicted number is computed using a quasar surface density deduced from several studies of quasars in the field and using an estimate of the number of regions sampled. Both these numbers are very uncertain (see ref. 15), and the whole analysis has been questioned³³.

If one takes from Table 2 of Burbidge the observed (N_0) and his estimate of the expected ($\langle n \rangle_0$) number of quasars within some angular separation ($\theta_{\max}$) then one can set $\langle n \rangle_\Omega = N_0$ in equation (7) and compute the mean enhancement factor $\bar{\zeta}_B(\theta_{\max})$ required if gravitational focusing is to account for the entire excess, $\bar{\zeta}_B = (N_0/\langle n \rangle_0 - 1)/2$. Table 1 lists the data both for the full quasar-galaxy sample and for the subset discovered by Arp. Note that the table entries are cumulative both across and down and so are not statistically independent.

Burbidge's computation of $\langle n \rangle_0$ is probably more applicable to the Arp subset¹, so it is the values of $\bar{\zeta}_B$ from the upper portion

Table 1 Mean quasar density enhancement for Burbidge's compilation

Mag	N_0	$\theta_{\max} = 1'$ $\langle n \rangle_0$	$\bar{\zeta}_B$	$\theta_{\max} = 2'$ $\langle n \rangle_0$	$\bar{\zeta}_B$	$\theta_{\max} = 3'$ $\langle n \rangle_0$	$\bar{\zeta}_B$
Arp sample							
<18	0	0.17	—	4	0.7	2.3	5
<19	1	0.52	0.46	10	2.1	1.9	12
<20	3	1.73	0.37	15	6.9	0.6	19
Total sample							
<18	5	0.17	14.0	11	0.7	7.3	12
<19	9	0.52	8.2	21	2.1	4.5	25
<20	12	1.73	3.0	28	6.9	1.5	34

N_0 is the observed number of quasars and $\langle n \rangle_0$ the predicted number (for Arp sample—see text) from Burbidge¹.

of Table 1 that might be compared with equation (8). These $\bar{\zeta}_B$ s exceed our nominal estimate of ζ by factors of three or more. Unfortunately, these $\bar{\zeta}_B$ s have little or no significance because of their large statistical errors (most of the contribution to the $\bar{\zeta}_B$ s comes from ~ 8 too many quasars brighter than 19 mag with $1' < \theta_{\max} < 2'$) combined with their sensitivity to unknown selection effects. Further data would be required to establish a real discrepancy.

The very large values of $\bar{\zeta}_B$ computed for the total quasar-galaxy sample cannot be taken seriously because the predicted $\langle n \rangle_0$ s are surely underestimated for this case. However, it is interesting that the relative values of $\bar{\zeta}_B$ in both portions of Table 1 are qualitatively in agreement with equation (8): for a given limiting magnitude, $\bar{\zeta}_B$ does in fact go roughly as $\sim \theta_{\max}^{-1}$; for a given $\theta_{\max}$, $\bar{\zeta}_B$ does fall as the limiting magnitude increases from 18 to 19 in approximate accord with the fall in F of Fig. 1. It is unlikely that the quasar sample is complete to 20 mag, so it is not too surprising that the same scaling fails when one compares $\bar{\zeta}_B$ s for magnitudes 19 and 20.

Burbidge also notes that the quasar-galaxy angular separation is inversely correlated with the galaxy redshift. This is consistent with the hypothesis of enhancement by gravitational focusing because for fixed R , the angular size of the region of enhanced quasar surface density would fall with increasing galaxy redshift.

Burbidge¹ and Arp² list several special cases of multiple quasar-galaxy associations. The calculation of the statistical significance of these cases is very difficult because of the unknown sample size and selection criteria. Furthermore, their significance is very sensitive to the assumed quasar surface density because one always has $\langle n \rangle_0 \ll 1$ for which the Poisson probability of detecting i quasars goes as $\langle n \rangle_0^i$. Of course, this also means that for quasars observed through the halo of a galaxy, the probability goes as $\langle n \rangle_0^i \sim \langle n \rangle_0^i (1 + 2\zeta(\theta))^i$ (from equation (7)), so that even modest values of ζ could have a significant effect on the inferred significance.

There are several special cases for which a density enhancement is not a sufficient explanation. These include quasar pairs with similar redshifts, which imply clustering or super-clustering (see refs 15, 34, 35) and quasars with supposedly unlikely orientations or alignments^{1,2} (but see ref. 36).

Conclusions

I have shown that the quasar surface density can be significantly enhanced by the chance alignment of faint quasars with individual stars in the haloes of foreground galaxies, stars that act as gravitational lenses to intensify the quasar light. The estimated magnitude of the effect, which has a large uncertainty due to our poor knowledge of the relevant parameters, is sufficient to explain fully the association of distant quasars with foreground galaxies found in the statistical analysis of Seldner and Peebles³. With nominal parameters the effect falls short of Burbidge's¹ estimates of the enhanced quasar density derived from Arp's quasar-galaxy associations. However, these estimated enhancements have little or no significance when statistical and systematic uncertainties are considered. Burbidge's estimates of quasar density enhancements near galaxies derived from his full catalogue do have approximately the dependences on angle and limiting magnitude that one would expect if the enhancements were due to gravitational focusing.

I conclude that the identification of this physically plausible mechanism for obtaining galaxy-quasar associations of the appropriate magnitude undermines the claims that such associations must be the symptoms of non-cosmological redshifts for some quasars.

There are several observations which could test the significance of gravitational focusing by stars on the quasar surface density. Clearly, an increased quasar sample with well defined statistical properties is extremely important. The growing sample of X-ray selected quasars from the Einstein Observatory may constitute such a set. Tests of this sample by the techniques of Seldner and Peebles³ would have to be consistent

with the numerical estimates of ζ computed with what one hopes will be continually improved galaxy parameters. Such an exercise could establish the existence of massive galaxy haloes composed of macroscopic objects, an existence already suggested by the agreement with the Seldner and Peebles data.

In addition to statistical tests, one can search for peculiarities in individual objects indicative of gravitational focusing. First, consider long-term variability of the quasar caused by the relative motion of the Earth and the star acting as the gravitational lens. Stellar velocities are ~ 200 – 300 km s^{-1} , and the galaxy may have some transverse bulk velocity relative to Earth. The impact parameter b of an undeflected quasar light ray hitting Earth is between $b = b_0 \sim A_B \alpha_0 = 4(GM_* A_B/c^2)^{1/2}$, and $b \geq 0.05b_0$. If $M_* = 1 M_\odot$ and $A_B \sim 10 \text{ Mpc}$ then $b_0 \sim 9 \times 10^{15} \text{ cm}$, and one would expect the quasar and star to become misaligned in $\sim 30(b/b_0)(v/100 \text{ km s}^{-1})^{-1} \text{ yr}$, where v is the component of the stellar velocity perpendicular to the line-of-sight and in the Earth-star-quasar plane. The misalignment will cause the previously magnified quasar to fade once and for all typically by several magnitudes. If none of the Burbidge objects showed this behaviour over a ~ 20 -yr period my conclusion would be untenable (unless $M_* \gg 1 M_\odot$). Conversely, if it can be shown that intensification by gravitational focusing does occur for some of Burbidge's objects, that they persist for $> 1 \text{ yr}$ already implies that M_* is not $\ll 10^{-3} M_\odot$; with a significant sample of objects it might someday be possible to estimate a mean value for M_* (this suggestion is also made by Gott⁹).

Another possible test concerns the short-time scale variability of the quasar. There will be a time delay of $\Delta t \sim \alpha_0 b_0/c = 16GM_*/c^3$ between light signals arriving at Earth from a point in the quasar along the two possible trajectories that pass the star¹³. Any variations in the quasar signal, whether inherent or due to scintillations along the path from the quasar to the star, would cause an enhancement in the autocorrelation function of the signal at a lag time Δt . Unfortunately, for $M_* = 1 M_\odot$, $\Delta t \sim 80 \mu\text{s}$, so this is an exceedingly difficult measurement. Nevertheless, a positive detection would validate the hypothesis and measure M_* directly.

Finally, if the dimensions of the broad line emission region are $> l_0$, then the lines would generally be intensified less than the continuum. In this case, intensified quasars would have anomalously weak lines, but the lines would fade less slowly than the continuum as the quasar and star became misaligned. Any quasar showing narrow line emission is not being intensified unless $M_* \gg 1 M_\odot$.

I thank George Clark, Irwin Shapiro and Julian Krolik for helpful comments. Julian Krolik first suggested that intensified quasars might have weak emission lines. I received helpful comments on the initial manuscript from C. Arp, J. Bahcall, J. R. Gott and M. Schmidt. My attention was drawn to the question of quasar-galaxy associations during the 1980 Astrophysics Workshop at the Aspen Center for Physics; I thank the Center for its hospitality and NASA for its support of the workshop. I thank Brenda Parsons and Maggie Carracino for preparing the manuscript. This work was supported in part by NASA contract NAS-8-30752. I am an Alfred P. Sloan Foundation Fellow.

Received 10 March; accepted 13 May 1981.

- Burbidge, G. *Nature* **202**, 451–455 (1979).
- Arp, H. *Ann. N.Y. Acad. Sci.* **336**, 94–112 (1980).
- Seldner, M. & Peebles, P. J. E. *Astrophys. J.* **227**, 30–36 (1979).
- Roberts, D. & O'Dell, S. *Astr. Astrophys.* **76**, 254–256 (1979).
- Gott, J. R. & Gunn, J. E. *Astrophys. J. Lett.* **190**, L105–L108 (1974).
- Barnothy, J. M. & Barnothy, M. F. *Astrophys. J.* **174**, 477–482 (1972).
- Barnothy, J. *Bul. Am. astr. Soc.* **6**, 212 (1974).
- Chang, K. & Refsdal, S. *Nature* **282**, 561–564 (1979).
- Gott, J. R. *Astrophys. J.* **243**, 140–146 (1981).
- Turner, E. L. *Astrophys. J. Lett.* **242**, L135–L140 (1980).
- Einstein, A. *Science* **84**, 506–7 (1936).
- Refsdal, S. *Mon. Not. R. astr. Soc.* **128**, 295–306 (1964).
- Refsdal, S. *Mon. Not. R. astr. Soc.* **132**, 101–111 (1966).
- Young, P., Gunn, J. E., Kristian, J., Oke, J. & Westphal, J. *Astrophys. J.* **241**, 507–520 (1980).
- Wills, D. *Phys. Scr.* **17**, 333–337 (1978).
- Gilmore, G. *Mon. Not. R. astr. Soc.* **195**, 183–196 (1981).
- Green, R. F. & Schmidt, M. *Astrophys. J. Lett.* **220**, L1–L4 (1978).
- Bohuski, T. & Weedman, D. *Astrophys. J.* **231**, 653–658 (1979).

19. Bahcall, J. & Sonnerup, R. M. *Astrophys. J. Lett.* **238**, L17–L20 (1980).
20. Gott, J. R. *Ann. Rev. Astr. Astrophys.* **15**, 235–266 (1977).
21. Faber, S. M. & Gallagher, J. S. A. *Rev. Astr. Astrophys.* **17**, 135–188 (1979).
22. Whitmore, B., Kirshner, R. & Schechter, P. *Astrophys. J.* **234**, 68–75 (1979).
23. Faber, S. M. & Jackson, R. E. *Astrophys. J.* **204**, 668–683 (1976).
24. Dressler, A. *Astrophys. J.* **231**, 659–670 (1979).
25. Peterson, C. J. & King, I. R. *Astr. J.* **80**, 427–436 (1975).
26. Davidson, K. & Netzer, H. *Rev. Mod. Phys.* **51**, 715–766 (1979).
27. Press, W. H. & Gunn, J. E. *Astrophys. J.* **185**, 397–412 (1973).
28. Weinberg, S. *Gravitation and Cosmology*, 191 (Wiley, New York, 1972).
29. Groth, E. J. & Peebles, P. J. E. *Astrophys. J.* **217**, 385–405 (1977).
30. Seldner, M., Siebers, B., Groth, E. & Peebles, P. J. E. *Astr. J.* **82**, 249–256 (1977).
31. Schechter, P. *Astrophys. J.* **203**, 297–306 (1976).
32. Holmberg, E. *Galaxies and the Universe* (eds Sandage, A. & Kristian, J.) 154–155 (University of Chicago Press, 1975).
33. Weedman, D. W. *Astrophys. J.* **237**, 326–330 (1980).
34. Burbidge, E. M., Juncarinen, V. T., Koski, A. T., Smith, H. E. & Hoag, A. A. *Astrophys. J. Lett.* **242**, L55–L59 (1980).
35. Oort, J. H., Arp, H. & de Ruiter, H. Preprint No. 129 (ESO, 1980).
36. Edmunds, M. G. & George, G. H. *Nature* **290**, 481–483 (1981).

Sunspot periods in the late Precambrian glacial climate and solar–planetary relations

G. E. Williams

Exploration Department, The Broken Hill Proprietary Company Limited, PO Box 559, Camberwell, Victoria 3124, Australia

Cycles conspicuous in late Precambrian (~680 Myr) glacial varves in Australia reveal strong climatic periods of about 11, 22, 145 and 290 yr and a weaker period near 90 yr. The 11-, 22- and 90-yr cycles equate with sunspot periods, and the longer cycles with solar and climatic periods as indicated by tree-ring studies. The varve cycles provide clear evidence of 11- and 22-yr climatic periods in the geological record and may carry important implications for solar physics and planetary science.

THE relation between sunspot cycles and patterns of terrestrial weather and climate remains controversial^{1–3}. The 11- and/or 22-yr solar periods have long been perceived in numerous meteorological and climatic variables^{2,4,5}, and recent studies have identified a 22-yr period in tree rings^{6,7}. Others, however, suggested^{8–11} that little or no convincing evidence exists for statistically significant correlations of either the 11- or 22-yr sunspot cycle with variations of weather and climate.

With regard to the geological record, periods of ~11- and ~22-yr and more, perceived in certain 'varved' (annually layered) sequences principally of Phanerozoic age, have been ascribed a solar origin^{12–16}. With rare exception¹⁶, however, such cyclicity is not conspicuous but is revealed only through the statistical processing of varve thickness. In the absence of a convincing explanation of solar-cycle influence on weather or of clear empirical evidence for such cyclicity in the climatological or geological record, strong doubts remain about the reputed terrestrial effects of sunspot variability.

Late Precambrian varve cycles described here provide the best evidence yet of 11- and 22-yr climatic periods in the remote past. The presence of such periods in rocks hundreds of millions of years old may have important implications for solar physics and planetary science.

Late Precambrian cyclic lamination

A cyclic lamination is conspicuous in the late Precambrian Elatina Formation at Pichi Richi Pass [32°25' S, 137°59' E] and at other localities up to ~30 km distant in the Flinders Ranges, South Australia. There the formation comprises mainly red and brown siltstones and fine sandstones with a thickness of ~30–60 m (refs 17–19). The Elatina Formation is a distal facies of western provenance deposited on a western 'shelf' of the Adelaide Geosyncline during the Marinoan glaciation^{17,19}. Correlative proximal outwash sandstones occur on the shelf ~100–150 km north-west of Pichi Richi Pass, indicating the former presence of an ice sheet on the low-relief cratonic region west of the shelf. Marinoan glaciation occurred ~680 Myr ago²⁰ in low palaeolatitudes²¹.

The cyclic lamination (Fig. 1) is characterized by the regular repetition of dark red brown, clayey bands spaced 2–16 mm apart that bound groups of paler, clastic laminae. The dark bands typically are 0.1–0.5 mm thick and the clastic laminae 0.1–2.5 mm thick. The lamination is essentially planar and

laterally persistent over several hundred metres within the limits of outcrop.

The clastic laminae typically comprise a lower, paler, sandy layer and an upper, silty layer. The contact between such layers is usually gradational, whereas contacts between adjacent laminae are relatively sharp. Microscopically, the clastic laminae consist mainly of grains of quartz (dominant) and feldspar in a clayey matrix. Grading results from an upward reduction in grain size from 20–100 µm (silt to very fine sand) at the base to 20–50 µm (silt) at the top of each lamina and an accompanying upwards increase in the proportion of matrix. A few laminae (usually the thickest lamina in a particular group) contain one or two thin, impersistent layers of silt or very fine sand. Such layers are regarded as internal features, rather than discrete clastic laminae, because of their impersistence and thinness compared with adjacent laminae. The dark bands comprise a fine-grained mixture of clayey material and iron oxide. Some bands that appear paler are seen under magnification to consist of several very thin (≤0.2 mm) clayey layers interleaved with silty laminae. Such a composite dark band occurs between cycles 9 and 10 in Fig. 1a.

The inspection of outcrops and sawcuts suggested that a uniform number of clastic laminae (about 10) occur in each group or 'cycle' between successive dark bands. Thin sections (Fig. 1a), however, reveal additional, very thin (≤0.2 mm) laminae of silt at the bottom and/or top of most cycles. Sixteen successive relatively thick cycles so examined contain 9–14 discrete clastic laminae, and average 11.3 ± 1.6 (Table 1). Two additional sequences each of 10 successive cycles studied in thin section average 11.2 ± 1.7 (range 9–14) and 11.0 ± 1.0 (range 9–12) clastic laminae per cycle. Although counts for several cycles have uncertainties of ± 1 because of the indistinctness of some laminae, the results overall are consistent and demonstrate the regularity of the cyclicity. The mean for the 36 cycles so examined is 11.2 laminae, and this value is used here to estimate laminae numbers for long sequences. Thin cycles from other stratigraphic levels of the Elatina Formation (Fig. 1b, c) seem to contain about the same number of clastic laminae as the thick cycles, but the thinness of laminae hinders accurate counting.

The thickness of clastic laminae varies systematically, attaining a maximum near the centre of each cycle (Figs 1a and 2). Most cycles show some asymmetry, as the thickest lamina usually occurs below or above the lamina of median number. Of the cycles listed in Table 1, nine show positive skewness (for

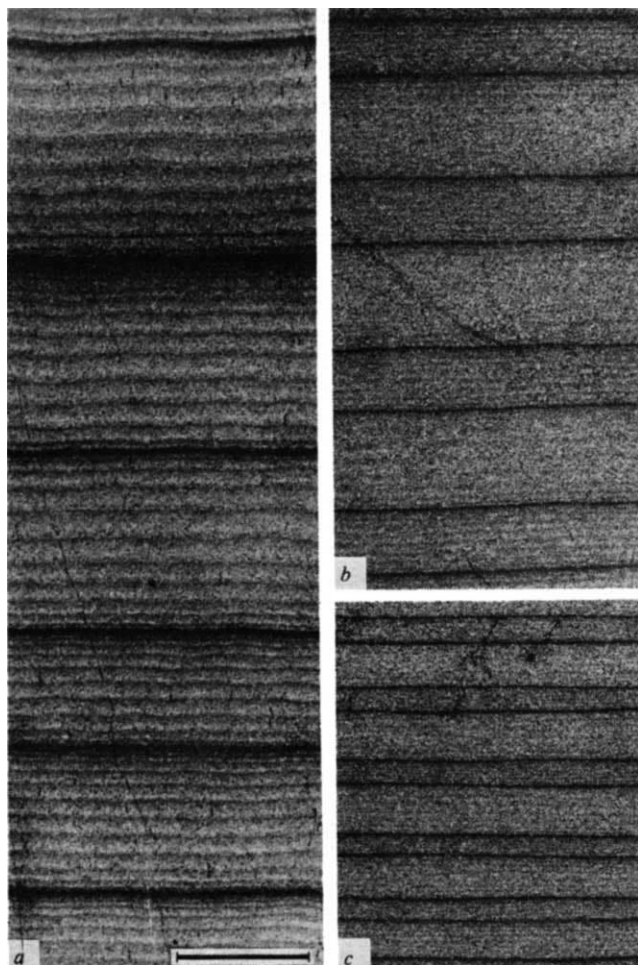


Fig. 1 Cyclic lamination of the Elatina Formation (late Precambrian) at Pichi Richi Pass, South Australia. Dark clayey bands spaced 2–16 mm apart bound cycles of pale, graded laminae of silt to very fine sand grade. *a*, Cycles 9 (base) to 13 (top) of the 16 cycles listed in Table 1. *b* and *c*, 'double cycles' resulting from alternate thick and thin successive cycles. Scale bar, 1 cm.

example, Fig. 2*a, b, d*) and five exhibit negative skewness (Fig. 2*c*). Overall, the 16 cycles tend to positive skewness (Table 1, column 3).

A 'double cycle' of alternate thick and thin cycles is common (Fig. 1*b, c*). Such double cycles are most conspicuous where individual cycles are relatively thin, making accurate counting of clastic laminae difficult. From the laminae counts for thicker cycles, which also exhibit double cycles, it may be concluded, however, that double cycles each contain on average 22–23 clastic laminae. The relative paleness of certain dark bands produces another type of double cycle (for example, cycles 9 and 10 in Fig. 1*a*, which together contain 23 clastic laminae).

Sequences of cycle thicknesses measured between successive dark bands seen on joint-faces and sawcuts perpendicular to the bedding reveal relatively long-term periodicity. Two consecutive sequences containing a total of 157 cycles ($\times 11.2 \approx 1,760$ clastic laminae) display several regularities (see Fig. 3):

(1) The double cycle is prominent through the alternation of thick and thin successive cycles and, where such a pattern may be absent, the presence of paler dark bands between alternate cycles.

(2) Conspicuous long-term variations in cycle thickness are similar in the two sequences. Significant minima occur every 12–14 cycles (~ 135 – 155 clastic laminae) and deepest minima every 25–27 cycles (~ 280 – 300 laminae); significant maxima

occur at intervals of 7–19 cycles (~ 80 – 215 laminae) and greatest maxima every 26–27 cycles (~ 290 – 300 laminae). The overall pattern suggests two dominant mean periods, one near 26 cycles (~ 290 laminae) and the other of 13 cycles (~ 145 laminae). The two periods are in the ratio 2:1 and hence represent a fundamental and a second harmonic. The periods of ~ 145 and ~ 290 laminae are conspicuous at other stratigraphic levels and seem to characterize the laminated facies of the Elatina Formation.

(3) Paler dark bands between alternate cycles, best seen on sawcuts, occur in similarly-placed groups within three successive major oscillations (Fig. 3*a*). The stratigraphic occurrence of such paler bands is apparently an integral part of the long-term periodicity identified.

Spectral analysis confirmed the visually-evident periods and revealed also a weaker period of ~ 90 laminae. Fourier transform of the precisely measured sequence (Fig. 3*a*) identified statistically significant periods (those that can oscillate about 10 times or more in a data sequence²²) of 22.4 laminae (relative power 2.51) and 89.6 laminae (0.44). For the full sequence (Fig. 3*a, b* combined) significant periods identified by Fourier transform are 22.4 laminae (4.89), 89.6 laminae (0.51) and 145.6 laminae (1.37). A period near 285 laminae (1.8) present in the spectrum for the full sequence is in close agreement with the visually-evident period of ~ 290 laminae. Fourier series analysis of the same sequences confirmed the presence of strong periods near 145 laminae (mean relative power 2.46) and near 290 laminae (4.12).

Late Precambrian climatic cycles

Sedimentary features exhibited by the Elatina Formation—cyclic graded lamination, sporadic graded beds several centimetres thick, cross-laminated sandstones, and rare glacial erratics^{23,24}—together suggest deposition mainly by density- and traction-currents in a proglacial lake.

Density currents are of great importance in glaciolacustrine sedimentation^{25–29}. During the spring and summer thaw, glacial meltwaters carry abundant suspended matter across the outwash plain to the proglacial lake. There the cold, turbid meltwaters form a density underflow which spreads a thin, commonly graded bed of silt and fine sand widely over the lake floor. Suspended clayey sediment may later settle from quiet waters during the winter when the lake may freeze over, or

Table 1 Data for 16 successive cycles of the Elatina Formation (cycles defined by dark bands)

Cycle	No. of clastic laminae*	Position of thickest lamina from cycle base*
16 (top)	12	4
15	10	6
14	9	4
13	11	6–7
12	14	5
11	11	7
10	11	4
9	12	5
8	13	4
7	14	4
6	11	5
5	12	6
4	12	5
3	9	5–6
2	9	5
1 (base)	10	4–5
Total	180	
Mean ($\pm$ s.d.)	11.3 ± 1.6	5.0 ± 1.0

* Data obtained through the photographic enlargement and microscopic study of thin sections.

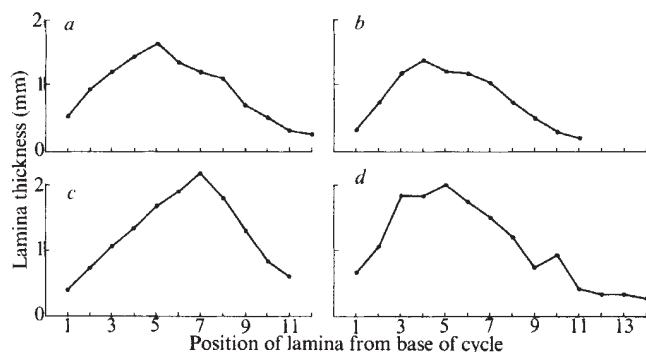


Fig. 2 Variation in thickness of clastic laminae within four cycles shown in Fig. 1a. a, Cycle 9; b, cycle 10; c, cycle 11; d, cycle 12. Thicknesses of laminae were measured to an accuracy of ~5–10% from 3 to 5 times photographic enlargements of thin sections.

through overturn of lake waters. Clayey winter layers are not always present, however, as summer (seasonal) influx of suspended matter into an oligotrophic lake whose waters are unstratified during influx or do not undergo seasonal overturn can produce graded laminae lacking a clayey uppermost layer²⁹. Where such graded beds or coarse-fine couplets form a regular or rhythmic sequence they are usually interpreted as annual deposits, or varves (*sensu lato*)^{30,31}. Spasmodic summer storms and dam bursts may also generate stream flows and density currents, resulting in the deposition of several random layers for that year. Glacial varves range in thickness from tens of centimetres to <1 mm, some being too thin to be measured³¹. The

thickness of clastic (summer) varves at a particular site may reflect the volume of meltwaters and their suspended load entering the lake; thick clastic varves therefore commonly correspond to warm summers and thin clastic varves to cold summers³².

The clastic laminae of the Elatina Formation accordingly are ascribed to deposition by density currents probably in a proglacial lake whose waters normally were not subject to winter freezing or overturn. The non-random variation in thickness of such laminae (Fig. 2) indicates a non-random fluctuation also in the volume of successive meltwaters and their suspended load entering the lake. Such regularity of hydrologic and sedimentary processes argues forcibly that the clastic laminae are not the result of random density currents but are indeed annual increments, or varves. Internal layers attributable to spasmodic density currents are uncommon.

As the thickness of clastic varves at a particular site may record relative warmth of summers³², the regular variation in thickness of varves of the Elatina Formation suggests the cyclic waxing and waning of mean summer temperatures with a mean period near 11 yr. It follows that the dark clayey bands bounding each cycle of ~11 summer laminae were deposited during the coldest portion of the climatic spectrum. The dark bands therefore may be attributed to the deposition of suspended fines during severe winters when the lake froze over or major overturn of cold surface waters occurred. The paler bands suggest a succession of less severe winters at the close of a cycle with limited freezing or overturn. A winter origin for the dark bands implies that the duration of each cycle is recorded by the number of summer clastic laminae.

I conclude therefore that the laminae of the Elatina Formation record a remarkably cyclic climate, with a basic period near

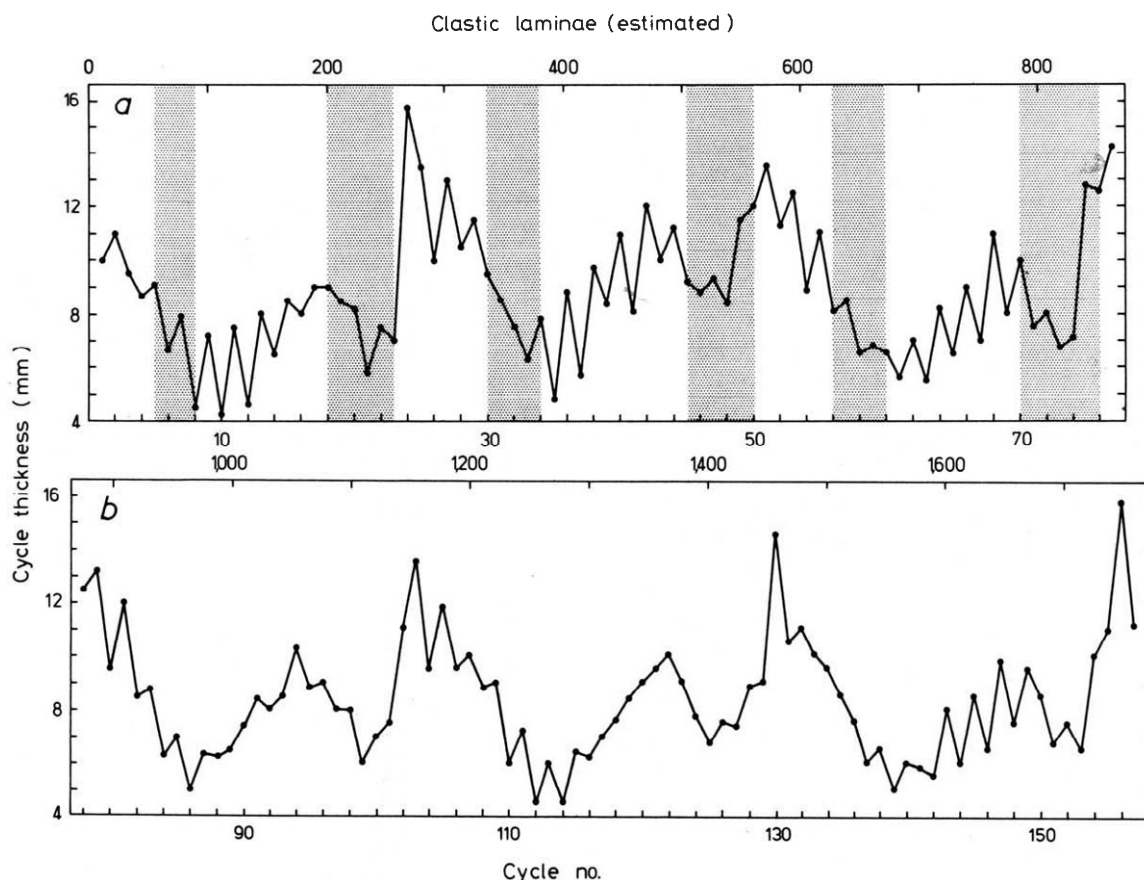


Fig. 3 Long-term trends in cycle thickness for two consecutive vertical sections measured between successive dark bands in the Elatina Formation. The mean cycle is taken as 11.2 clastic laminae. Cycle numbers shown increase up-sequence. a, Sawcut; shaded intervals indicate paler dark bands between alternate cycles. b, Measured in the field; section immediately overlies that shown in a.

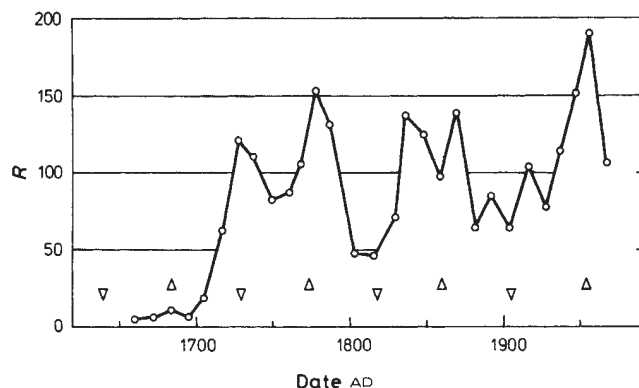


Fig. 4 Annual mean sunspot number R at maxima of the 11-yr solar cycle, from AD 1645 to the present. Triangles mark extrema of the 80–90-yr 'Gleissberg cycle' (from ref. 58).

11 yr, which prevailed in southern Australia during the late Precambrian Marinoan glaciation. During such cycles, summers gradually became warmer over the first 4–7 yr with accompanying increase in the volume and suspended load of glacial meltwaters. The climate then progressively deteriorated over 4–10 yr and the volume and load of meltwaters decreased until at the close of each cycle frigid conditions prevailed. The common alternation in cycle thickness to give a period of ~22 yr (Fig. 1*b, c*) suggests the attainment of alternately high and low summer temperatures during successive cycles. Moreover, the interpretation of varve and ~11-yr cycle thicknesses as measures of relative summer temperature implies that long-term variations in cycle thickness (Fig. 3) in effect record long-term fluctuations in relative summer temperature. Hence long-term climatic periods near 90, 145 and 290 yr are revealed, with significant temperature-minima occurring every 135–155 yr and deepest minima every 280–300 yr. Less severe winter end-cycle conditions, as recorded by paler dark bands, appear confined to times of lesser temperature-minima and to intervals immediately preceding the deepest temperature-minima (Fig. 3*a*). Importantly, the climatic-temperature fluctuations were of regional significance, as they apparently affected both the summer meltwater discharges from a continental-type ice sheet and the winter freezing or overturn of an extensive proglacial lake situated ~100 km to the east.

Cause of the climatic cyclicity

The climatic cycles identified may reflect an endogenetic mechanism such as glacial surging, or an exogenetic, solar influence. The idea of climatic control by glacial surging encounters major difficulties:

(1) Surging glaciers^{33–35} undergo sudden advance after years of stagnation or very slow retreat. Medvezhii in the Pamirs, for example, has a surge period of 10–14 yr and a surge duration of only 0.2 yr (ref. 33), and mountain glaciers in western North America that have surge periods of 19–20 yr have active phases of only 2–3 yr (ref. 34). The ~11-yr cycles of the Elatina Formation indicate, in contrast, a gradual increase in meltwater discharge over 4–7 yr followed by a waning of such discharge over 4–10 yr. The structure of the ~22-yr double cycles—alternate thick and thin (warmer and cooler) ~11-yr cycles—also seems unrelated to the pattern of glacial surging.

(2) Surging glaciers that have short surge-periods of ≤ 20 yr typically are relatively small, steep mountain-glaciers that fall through ~1,000–2,000 m and undergo maximum advances of only ~1–5 km (refs 33, 34). In contrast, the regional palaeogeography of the Elatina Formation and correlative shelf units indicates a continental-type ice sheet on a low-relief cratonic area discharging meltwaters to a lake some 100 km distant.

(3) The cycles in the Elatina Formation suggest regional climatic-temperature changes that affected both the ice sheet and lake. A small glacier that periodically advanced only a few kilometres is an unlikely cause of such widespread climatic changes. Indeed, glacial surges are unrelated to climatic fluctuations^{34,35}.

(4) A surging glacier maintains a reasonably constant surge period^{34,35}. The varve-cycles of the Elatina Formation, in contrast, exhibit several strong periods of between ~11 and ~290 yr.

The surging of a glacier therefore seems an improbable control of the climatic cyclicity recorded by the Elatina Formation. Hence I shall examine the Elatina record in detail for evidence of possible solar influence.

Table 2 compares data for 16 successive varve-cycles with two centuries of sunspot data. General correspondence exists between all respective means and ranges for the two sets of data. Moreover, the 16 cycles, which were chosen for study because of their large thickness, tend to positive skewness with respect to varve thickness (Table 1). Interestingly, sunspot cycles of large sunspot maxima on average show positive skewness, the time of rise to maximum sunspot number generally being less than the time of fall to sunspot minimum³⁶. Another relevant feature of the Elatina record is the common alternation of thick and thin ~11-yr cycles to produce ~22-yr double cycles. As noted^{36,37}, sunspot activity between AD 1840 and 1940 exhibited a systematic alternation of high and low maxima for successive 11-yr cycles; the 'zig-zag' pattern of sunspot maxima (Fig. 4) finds similarity in plots of varve-cycle thickness (Fig. 3).

Modern analytical techniques have identified a period near 90 yr in sunspot data since about AD 1700 (refs 2, 38–40), thus confirming the ~80–90 yr 'Gleissberg cycle' (Fig. 4). The period of ~90 yr indicated by Fourier transform of Elatina Formation varve-cycle thicknesses therefore accords with an established sunspot period, but the two strong periods of ~145 and ~290 yr find no apparent counterparts in the observed solar record. The available record of the sunspot number covers less than 300 yr, however, and proposed periods exceeding, say, 90 yr must be regarded with utmost caution⁴⁰. In comparison, the long-term changes in ~11-yr varve-cycle thickness (Fig. 3) cover a total interval of ~1,760 yr.

Study of ¹⁴C levels in tree rings suggests four important minima in solar activity since about AD 1000 (Maunder, AD 1654–1714; Spörer, 1416–1534; Wolf, 1282–1342; and a minimum near AD 1040), whose midpoints are spaced respectively 209, 163 and 272 yr apart⁴¹. The spacing of these inferred solar minima, although irregular, is of the same order as that suggested here for late Precambrian temperature-minima (135–155 and 280–300 yr). Moreover, one or both of the mean

Table 2 Comparison of data for varve cycles in the Elatina Formation and for recent sunspot cycles

	Mean ±s.d. (yr)	Range (yr)
Data for 16 successive Elatina Formation cycles defined by dark bands (180 yr) (see Table 1)		
Cycle period	11.3 ± 1.6	9–14
Period between thickest varves in successive cycles	11.2 ± 2.4	8–15
Time of rise to maximum varve-thickness in cycle	5.0 ± 1.0	4–7
Time of fall to end of cycle	6.3 ± 2.0	4–10
Sunspot data for 203 yr (19 successive cycles, AD 1755–1958) (ref. 59, p 18):		
Period between minima	11.1 ± 1.3	9.0–13.6
Period between maxima	10.9 ± 2.3	7.3–17.1
Time of rise to maximum	4.5 ± 1.2	2.9–6.9
Time of fall to minimum	6.6 ± 1.4	4.0–10.2

climatic periods of ~145 and ~290 yr recorded by the Elatina Formation accord with: (1) the two dominant climatic periods of ~155 and 280 ± 10 yr found in oxygen and hydrogen stable-isotope ratios measured in 1,800 yr of Japanese cedar rings⁴²; (2) the double period of 284–303 yr and apparent second harmonic of 142–151 yr determined in tree-growth from 274 BC to AD 1914 (ref. 43); (3) the period of 290 yr obtained for δD values of bristlecone-pine rings between AD 970 and 1970 (ref. 7); and (4) the period of 286 yr found through ¹⁴C measurements on tree rings between 2,000 and 6,000 yr BP (ref. 44).

The ~11-, ~22- and ~90-yr climatic cycles recorded by the Elatina Formation are, therefore, in good agreement with modern sunspot periods, and the cycles of ~145 and ~290 yr with long-term solar and climatic periods as indicated by tree-ring studies. Accordingly, there may be a connection between varve cyclicity and solar variability. Accepting such a relation, the data in Table 2 imply that varve thickness, and therefore meltwater discharge, varied directly with sunspot activity. Maximum summer temperatures thus occurred at sunspot maxima, and the coldest conditions at sunspot minima. By this interpretation, Fig. 3 illustrates long-term trends not only in relative summer temperature but also in relative maximum sunspot number. So delicately balanced were the climatic conditions that only at or near sunspot minima did winter freezing or major overturn of lake waters occur, thus giving the climatic cyclicity such a clear imprint in the sedimentary record.

Discussion

If the climatic cyclicity identified indeed reflects solar variability, then solar control of climate must have been much greater during deposition of the Elatina Formation than is evident today. The intensity of the geomagnetic field may modulate weather and climate^{45–47}, and palaeointensity variation plausibly may account for such apparent enhanced solar control of climate in the geological past.

The incidence of solar flares and the overall level of flare particle emission, comprising mainly protons, vary directly with the 11-yr solar cycle^{2,44,48}. Such corpuscular radiation is normally deflected at the bow shock of the magnetosphere, but some eventually may reach the Earth's atmosphere near the

magnetic poles^{49,50}. If the dipole component of the geomagnetic field were to weaken or disappear over a few thousand years during a polarity reversal, the shielding effect of the magnetosphere would be much reduced or eliminated; solar corpuscular radiation could then penetrate the atmosphere to lower levels and low magnetic-latitudes and blanket the Earth completely^{51–53}. At such times any influence that the sunspot cycle may exert on weather and climate through solar corpuscular radiation² would be greatly magnified. The clarity of the varve cycles of the Elatina Formation therefore may be due ultimately to a weakening of the geomagnetic field during a polarity reversal. An alternative explanation is provided by the observation that during the latest Precambrian and early Palaeozoic ~700–400 Myr ago the geomagnetic field apparently was much weaker than that of today^{54,55}; a palaeointensity near 10% of the present value is suggested at about the time the Elatina Formation was deposited⁵⁴. This seeming weakness of the late Precambrian geomagnetic field, if confirmed by further studies, might explain the apparently enhanced solar influence on climate at that time.

The empirical evidence presented here suggests that the 11- and 22-yr and longer solar cycles are stable features of the Sun over hundreds of millions of years. Such apparent very long-term stability of solar processes carries important implications for solar physics and models of the solar activity cycle⁵⁶. Moreover, it may allow the prediction of future solar (and, perhaps, climatic) variations from the detailed study of varve cycles where a Sun-climate connection can be shown. The apparent constancy of solar periods over geologically significant time-intervals also has important implications for the long-term constancy of planetary orbital periods¹⁶ and of the gravitational parameter G (ref. 57, p. 341).

The apparent imprint of sunspot periods in Precambrian varved rocks therefore has relevance for solar physics and several areas of planetary science. Ancient annually-layered rocks accordingly should be seen as potential solar observatories that may extend our understanding of solar processes and solar-planetary relations.

I thank P. G. Haslett, M. Kenyon and O. W. Polatayko for valuable assistance, and A. B. Pittock, C. M. Stacey and A. D. Stewart for helpful comments.

Received 22 December 1980; accepted 16 April 1981.

- Meadows, A. J. *Nature* **256**, 95–97 (1975).
- Herman, J. R. & Goldberg, R. A. *Sun, Weather, and Climate* (NASA, Washington, 1978).
- Scherrer, P. H. *Rev. Geophys. Space Phys.* **17**, 724–731 (1979).
- Lamb, H. H. *Climate: Present, Past and Future* Vol. 1, 440–464 (Methuen, London, 1972).
- King, J. W. *Astronaut. Aeronaut.* **13**(4), 10–19 (1975).
- Mitchell, J. M. Jr, Stockton, C. W. & Meko, D. M. in *Solar-Terrestrial Influences on Weather and Climate* (eds McCormac, B. M. & Seliga, T. A.) 125–143 (Reidel, Dordrecht, 1978).
- Epstein, S. & Yapp, C. J. *Earth planet. Sci. Lett.* **30**, 252–261 (1976).
- Gerety, E. J., Wallace, J. M. & Zerefos, C. S. *J. Atmos. Sci.* **34**, 673–678 (1977).
- Mass, C. & Schneider, S. H. *J. Atmos. Sci.* **34**, 1995–2004 (1977).
- Pittock, A. B. *Rev. Geophys. Space Phys.* **16**, 400–420 (1978).
- Robock, A. J. *Atmos. Sci.* **35**, 1111–1122 (1978).
- Bradley, W. H. *Prof. Pap. U.S. geol. Surv.* **158**, 87–110 (1929).
- Anderson, R. Y. *Ann. N.Y. Acad. Sci.* **95**, 424–439 (1961).
- Richter-Bernburg, G. in *Problems in Palaeoclimatology* (ed. Nairn, A. E. M.) 510–521 (Interscience, London, 1964).
- Duff, P. McL. D., Hallam, A. & Walton, E. K. *Cyclic Sedimentation* (Elsevier, Amsterdam, 1967).
- Trendall, A. F. *Econ. Geol.* **68**, 1089–1097 (1973).
- Preiss, W. V. *Dept Mines Energy, S. Aust., Rep. Bk. 79/24* (1979); in *Precambrian Geology of the Southern Continents* (ed. Hunter, D. R.) (Elsevier, Amsterdam, in the press).
- Dalgarno, C. R., Johnson, J. E., Forbes, B. G. & Thomson, B. P. *Port Augusta 1:250,000 Geol. map sheet* (Geological Survey South Australia, Adelaide, 1968).
- Coats, R. P. *Dept Mines Energy, S. Aust., Rep. Bk. 79/138* (1979).
- Coats, R. P. & Preiss, W. V. *Precamb. Res.* **13**, 181–208 (1980).
- McWilliams, M. O. & McElhinny, M. W. *J. Geol.* **88**, 1–26 (1980).
- Blackman, R. B. & Tukey, J. W. *The Measurement of Power Spectra* (Dover, New York, 1958).
- Dalgarno, R. & Johnson, J. E. *Q. Geol. Not. geol. Surv. S. Aust.* **11**, 3–4 (1964).
- Thomson, B. P. in *Handbook of South Australian Geology* (ed. Parkin, L. W.) 49–83 (Geological Survey of South Australia, Adelaide, 1969).
- Embleton, C. & King, C. A. M. *Glacial Geomorphology* (Arnold, London, 1975).
- Jopling, A. V. & McDonald, B. C. (eds) *Glaciofluvial and Glaciolacustrine Sedimentation* (SEPM Spec. Publ. 23, Tulsa, Oklahoma, 1975).
- Friedman, G. M. & Sanders, J. E. *Principles of Sedimentology* (Wiley, New York, 1978).
- Schlüchter, C. (ed.) *Moraines and Varves* (Balkema, Rotterdam, 1979).
- Sturm, M. in *Moraines and Varves* (ed. Schlüchter, C.) 281–285 (Balkema, Rotterdam, 1979).
- Gary, M., McAfee, R. & Wolf, C. L. (eds) *Glossary of Geology* (American Geological Institute, Washington, 1972).
- Mörner, N.-A. in *The Encyclopedia of Sedimentology* (eds Fairbridge, R. W. & Bourgeois, J.) 841–843 (Dowden, Hutchinson & Ross, Stroudsburg, 1978).
- Schove, D. J. in *Moraines and Varves* (ed. Schlüchter, C.) 319–323 (Balkema, Rotterdam, 1979); *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **25**, 209–233 (1978).
- Budd, W. & McInnes, B. in *Climatic Change and Variability. A Southern Perspective* (eds Pittock, A. B., Frakes, L. A., Jessen, D., Peterson, J. A. & Zillman, J. W.) 228–234 (Cambridge University Press, 1978).
- Meier, M. F. & Post, A. *Can. J. Earth Sci.* **6**, 807–817 (1969).
- Meier, M. F. *Can. J. Earth Sci.* **6**, 987–989 (1969).
- Bray, R. J. & Loughhead, R. E. *Sunspots* (Chapman & Hall, London, 1964).
- Dodson, H. W. & Hedeman, E. R. in *Solar-Terrestrial Physics 1970: Part I, The Sun* (ed. de Jager, C.) 151–172 (Reidel, Dordrecht, 1972).
- Cohen, T. J. & Lintz, P. R. *Nature* **250**, 398–400 (1974).
- Lomb, N. R. & Andersen, A. P. *Mon. Not. R. astr. Soc.* **190**, 723–732 (1980).
- Wittman, A. *Astr. Astrophys.* **66**, 93–97 (1978).
- Stuiver, M. & Quay, P. D. *Science* **207**, 11–19 (1980).
- Libby, L. M. & Pandolfi, L. J. *Nature* **266**, 415–417 (1977).
- Turner, H. H. *Mon. Not. R. astr. Soc.* **80**, 793–803 (1920).
- Damon, P. E. in *The Solar Output and its Variation* (ed. White, O. R.) 429–448 (Colorado Associated University Press, Boulder, 1977).
- Wollin, G., Ericson, D. B. & Ryan, W. B. F. *Earth planet. Sci. Lett.* **12**, 175–183 (1971).
- Wollin, G., Ryan, W. B. F. & Ericson, D. B. *Earth planet. Sci. Lett.* **41**, 395–397 (1978).
- King, J. W. *Nature* **247**, 131–134 (1974).
- Tandberg-Hanssen, E. *Solar Activity* (Blaisdell, Waltham, 1967).
- Kavanagh, L. D. Jr in *Earth's Magnetospheric Processes* (ed. McCormac, B. M.) 3–15 (Reidel, Dordrecht, 1972).
- Lanzert, L. J. in *The Solar Output and its Variation* (ed. White, O. R.) 383–403 (Colorado Associated University Press, Boulder, 1977).
- Siscoe, G. L. & Crooker, N. U. *J. Geomagn. Geoelectr.* **28**, 1–9 (1976).
- Siscoe, G. L., Chen, C.-K. & Harel, M. *J. Atmos. terr. Phys.* **38**, 1327–1331 (1976).
- Gribbin, J. in *Climatic Change* (ed. Gribbin, J.) 150–154 (Cambridge University Press, 1978).
- Carmichael, C. M. *Earth planet. Sci. Lett.* **3**, 351–354 (1967); in *Palaeogeophysics* (ed. Runcoorn, S. K.) 73–77 (Academic, London, 1970).
- McElhinny, M. W. *Palaeomagnetism and Plate Tectonics* (Cambridge University Press, 1973).
- Howard, R. & LaBonte, B. J. *Astrophys. J.* **239**, L33–L36 (1980).
- Williams, G. E. *Benchmark Papers in Geology* 57 (Hutchinson Ross, Stroudsburg, 1981).
- Eddy, J. A. in *The Solar Output and its Variation* (ed. White, O. R.) 51–71 (Colorado Associated University Press, Boulder, 1977).
- Waldmeier, M. *The Sunspot-Activity in the Years 1610–1960* (Schulthess, Zürich, 1961).

Hormonal inducibility of rat α_{2u} globulin genes in transfected mouse cells

David T. Kurtz

Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA

Two different genes coding for the hormonally regulated rat liver protein α_{2u} globulin were introduced into mouse Ltk⁻ cells through co-transfection with the HSV-1 thymidine kinase gene. Three to ten copies of the α_{2u} globulin genes were detected in several tk⁺ clones, over 50% of which could be induced with dexamethasone, to produce α_{2u} globulin mRNA and protein. This suggests that the information necessary for hormonal response is contained in the DNA fragment used for transfer.

THE ability to introduce defined DNA segments into mammalian cells¹ is potentially an extremely powerful tool for studying the regulation of eukaryotic gene expression, and for identifying specific DNA sequences which are responsible for such regulation. However, the usefulness of this technique depends critically on obtaining the correct, regulated expression of the foreign gene in the recipient cell. Although expression of rabbit globin mRNA² and ovalbumin³ in mouse L cells following gene transfer has been reported, in neither case was any evidence presented for regulation of expression.

α_{2u} Globulin is a male rat liver protein whose synthesis is under complex hormonal control *in vivo*: androgens, glucocorticoids and thyroid hormone induce α_{2u} globulin synthesis, and oestrogens strongly repress the synthesis of this protein^{4,5}. I have recently found that α_{2u} globulin is encoded by a multigene family (18–20 genes per haploid complement), clustered on chromosome 5 (ref. 6). Several α_{2u} globulin genes have been isolated from a library of the rat genome cloned in λ phage. Two of these genes were introduced into mouse Ltk⁻ cells (which are known to contain receptors for glucocorticoid hormones⁷), to determine if these rat genes retain hormonal inducibility in mouse cells in culture.

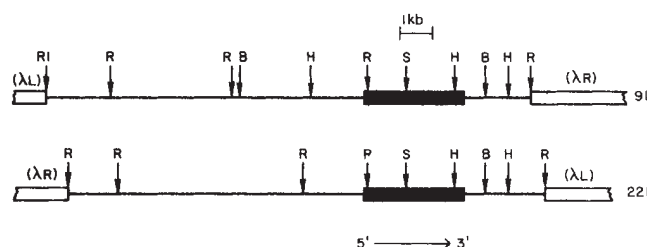


Fig. 1 The α_{2u} globulin genes co-transformed into mouse cells. Clones 91 and 221 are two of the nine α_{2u} globulin clones which have been isolated from the rat library. The coding regions (solid areas) are ~3.2 kilobases (kb) (compared to 1 kb for α_{2u} globulin mRNA) and each gene contains six introns (L. Chow and D.T.K. in preparation). Restriction enzyme sites indicated are: R, *EcoRI*; B, *BamHI*; S, *SalI*; H, *HindIII*.

Co-transformation of α_{2u} globulin genes into mouse Ltk⁻ cells

Nine different α_{2u} globulin genes (that is, genes contained on different-sized restriction fragments, and having distinct flanking sequences) have thus far been isolated from a library of the rat genome cloned in Charon 4A. Two of these genes (Fig. 1) were used for co-transformation into mouse Ltk⁻ cells, using a cloned HSV thymidine kinase gene as a vector⁸. Individual tk⁺ clones were isolated and grown in mass culture. High molecular weight DNA was extracted from the cell clones, cleaved with *BamHI*, fractionated on a 0.5% agarose gel and blotted on to

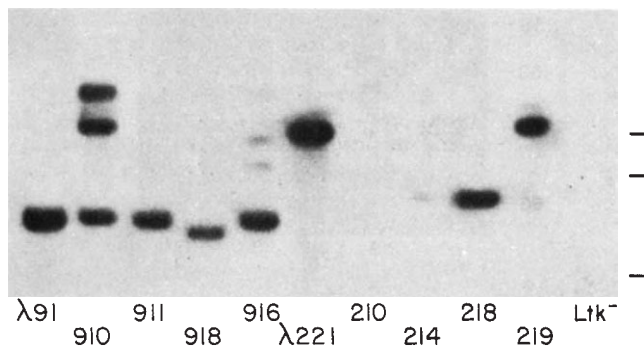


Fig. 2 Presence of α_{2u} globulin gene sequences in mouse L-cell clones. L cells were transformed using the calcium phosphate technique⁹ with 20 ng HSV tk gene, plus 500 ng of λ α_{2u} globulin recombinant plus 20 μ g carrier mouse Ltk⁻ DNA per 10^6 cells. Individual tk⁺ cell clones were isolated, DNA was extracted and cleaved with *BamHI*. 20 μ g were run on a 0.5% agarose gel, blotted on to nitrocellulose and probed with labelled α_{2u} globulin cDNA. Lanes marked ' λ 91' and ' λ 221' contain 500 pg of λ - α_{2u} recombinants cleaved with *BamHI*. (This corresponds to ~5 gene equivalents per 20 μ g cell DNA.) 910, 911, 916 and 918 are clones which are transformed with λ 91, and 210, 218 and 219 were transformed with clone 221.

nitrocellulose⁹. A cloned cDNA for α_{2u} globulin¹⁰ was labelled by nick translation and hybridized to the filters. It was found (Fig. 2) that the various Ltk⁺ clones had incorporated 3–10 copies of the α_{2u} globulin genes.

Introduction of α_{2u} globulin mRNA in response to dexamethasone

L-cell clones which were found to have incorporated the α_{2u} globulin genes were treated with dexamethasone (5×10^{-7} M) for 48 h. Messenger RNA was extracted from the induced and uninduced cells, fractionated on formaldehyde-agarose gels, blotted on to nitrocellulose¹¹ and probed with nick-translated α_{2u} globulin cDNA. α_{2u} Globulin mRNA was produced in these L cells in response to dexamethasone (Fig. 3). In several of the clones, a constitutive level of α_{2u} globulin mRNA was detectable even in the absence of dexamethasone and the mRNA level in the clones then rose in response to the hormone. No hybridization was found with mRNA from untransformed L cells with or without hormone. One clone (210) was found to produce a mRNA species in response to dexamethasone which hybridized to α_{2u} globulin cDNA, but was considerably larger than authentic α_{2u} globulin message. The nature of this heterogeneity is being investigated.

The induced level of α_{2u} globulin mRNA in these L-cell clones was determined by solution hybridization to single-stranded cDNA (Fig. 4). In a typical clone (918), the level was

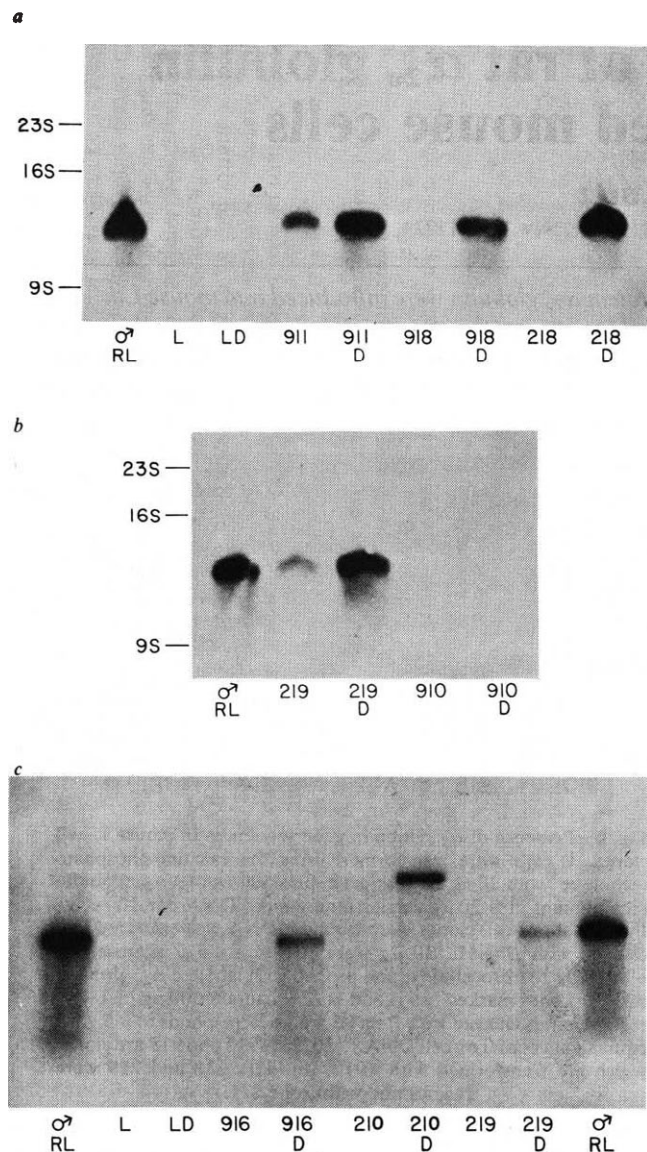


Fig. 3 α_{2u} Globulin mRNA produced in L cells in response to dexamethasone. L-cell clones containing the α_{2u} globulin genes were treated with 5×10^{-7} M dexamethasone for 48 h, and cytoplasmic mRNA was extracted as described previously¹⁰. The mRNA was fractionated on formaldehyde agarose gels, blotted on to nitrocellulose and probed with the labelled α_{2u} globulin cDNA clone. Lanes marked 'RL' contain 0.1 μ g of male rat liver mRNA; all other lanes contain 10 μ g of cell mRNA. (Lanes marked 'D' contain mRNA from hormonally induced cells). L, mRNA from untransformed Ltk⁻ cells. Note that clone 910 produces no detectable α_{2u} globulin mRNA, although it contains several copies of the gene.

~200 copies per cell, that is, 1/10th the level found in adult male rat liver⁷.

The cDNA for α_{2u} globulin shows slight cross-hybridization with the mRNA for MUP, the male mouse liver protein which corresponds to rat α_{2u} globulin^{10,12}. Therefore, it was important to confirm that the induced mRNA species in these mouse L-cell clones was indeed α_{2u} globulin, and not MUP mRNA. Hybrids formed between α_{2u} globulin cDNA and MUP mRNA melt at 69–70°C (Fig. 5), compared with a T_m of ~86°C for α_{2u} globulin mRNA. The T_m of the hybrid between α_{2u} globulin cDNA and the induced mRNA species in clone 918 was found to be identical to that for authentic rat liver α_{2u} globulin mRNA (Fig. 5).

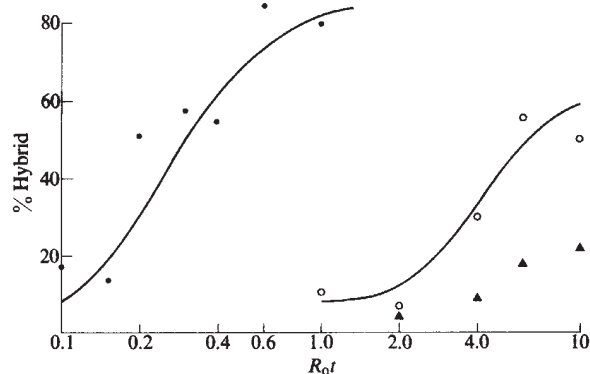


Fig. 4 The cDNA portion of p α 176, the pBR recombinant containing α_{2u} globulin cDNA¹⁰, was excised and nick translated (without DNase I) to a specific activity of 5×10^7 c.p.m. per μ g. The double-stranded cDNA (1 μ g) was denatured and hybridized to 500 μ g male rat liver mRNA in 80% formamide, 50 mM PIPES pH 7.4, 5 mM EDTA and 0.4 M NaCl at 55 °C for 2 h. RNA–DNA hybrids were isolated using hydroxylapatite chromatography⁴. The RNA was destroyed by treatment with alkali and the labelled single-stranded cDNA isolated after passage over Sephadex G-50. This single-stranded cDNA (average size ~300 bp) was annealed with male rat liver mRNA (●), or mRNA from transformed cell clone 918, control (▲) or dexamethasone-treated (○). Hybrid formation was determined using S_1 nuclease.

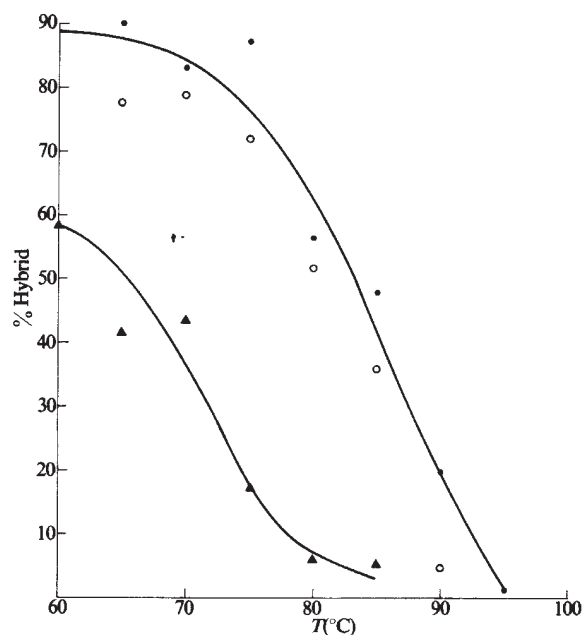


Fig. 5 Thermal denaturation of hybrids between α_{2u} globulin cDNA and MUP mRNA. Hybrids between α_{2u} globulin cDNA and rat liver (●), mouse liver (▲) or clone 918 (○) mRNA were allowed to form at 60 °C. The temperature was raised in 5 °C increments, the tubes were allowed to equilibrate for 10 min, and hybrids were assayed using S_1 nuclease.

Induction of α_{2u} globulin protein in L-cell clones

It was next of interest to determine if α_{2u} globulin protein was produced and secreted by the cells which contain α_{2u} globulin mRNA. The clones were treated with dexamethasone for 36 h and labelled with 35 S-methionine for 16 h. The medium was

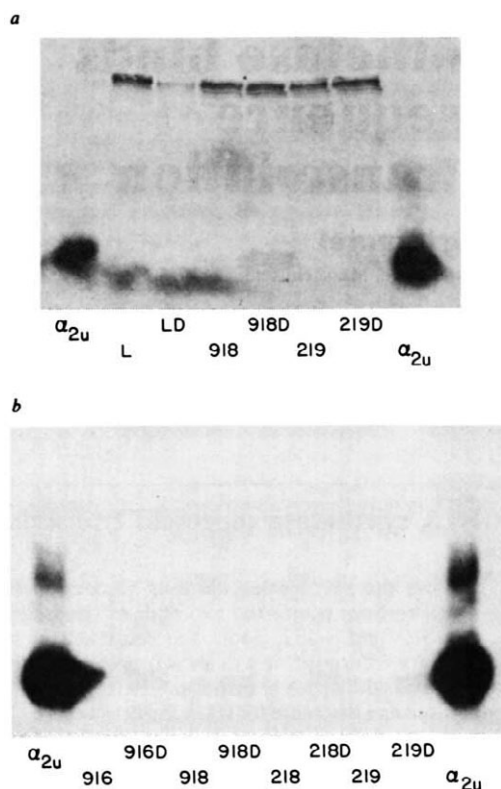


Fig. 6 α_{2u} Globulin protein immunoprecipitated from transformed cell media. Transformed cells were treated with dexamethasone for 36 h, labelled with ^{35}S -methionine (200 μCi per 10^7 cells) for 12 h. The media were collected, immunoprecipitated with rabbit anti- α_{2u} globulin and *S. aureus* (IgG Sorb, New England Enzyme Center). Boiled precipitates were subjected to SDS-gel electrophoresis on 20% polyacrylamide gels¹⁷, and the gels were dried and autoradiographed. Lanes marked ' α_{2u} ' contain purified α_{2u} globulin¹³ labelled with ^{125}I . Dexamethasone treatment did not enhance overall protein synthesis or secretion in these cells; in most clones, in fact, there was a slight decrease in ^{35}S -methionine incorporation in response to the hormone. *a* Represents a 3-day exposure, *b* a 7-day exposure.

collected and treated with monospecific anti- α_{2u} globulin¹³, followed by formalin-fixed *Staphylococcus aureus*. The bacteria were pelleted and washed thoroughly, and the bound protein was released⁵ and subjected to SDS-polyacrylamide gel electrophoresis (Fig. 6). ^{35}S -labelled α_{2u} globulin was found in the medium of the L-cell clones which produced α_{2u} globulin mRNA. Again, most of the clones showed a constitutive level of synthesis, which was then further induced in response to hormone. No immunoprecipitated material is found in control (untransformed Ltk⁻) cells.

Twelve L-cell clones were isolated which contained the entire α_{2u} coding region from λ clone 91; of these, eight were found to produce α_{2u} globulin mRNA and protein in response to

hormone. For λ clone 221, four of nine α_{2u} -containing L-cell clones were positive. There seems to be no correlation between α_{2u} gene copy number and the constitutive or induced level of α_{2u} globulin mRNA in the clones.

It is unclear whether the α_{2u} globulin mRNA is being transcribed from one or all of the genes which have been incorporated. Clone 210 produces a mRNA species which hybridizes to α_{2u} globulin cDNA but is ~600 base pairs (bp) larger than authentic α_{2u} globulin message. Further, this mRNA is the only ' α_{2u} globulin' mRNA produced by clone 210. This suggests that, for this particular clone, the transcription may indeed be occurring from only one of the genes, which perhaps has been incorporated into the carrier DNA adjacent to another promoter, or more likely, to another poly(A)-addition site. Preliminary evidence indicates that the large-sized α_{2u} globulin mRNA produced by clones 210 is not a result of incorrect splicing of a correctly transcribed α_{2u} globulin mRNA, because this mRNA does not hybridize to any intron-specific fragments of clone 221.

Discussion

I have shown that rat α_{2u} globulin genes, when introduced into mouse cells, respond to at least one of the hormones which normally induce α_{2u} globulin *in vivo*. (Ltk⁻ cells possess neither androgen nor oestrogen receptors, and thus, analogous studies concerning the response to these hormones must be performed in other tissue culture cells.)

There is no evidence that this hormonal induction of α_{2u} globulin in mouse cells is due to control of the transcription of these genes, although such a model is tempting to consider. Steroid hormones have profound effects on the cytoplasmic half-lifetimes of specific mRNAs in several systems^{14,15}, and thus the rise in the level of α_{2u} globulin mRNA in these mouse cells could be due to stabilization of the message. Studies are now being done to determine this possibility.

Nonetheless, whether the induction of α_{2u} globulin in these cells is due to transcriptional or post-transcriptional events, it should be possible to identify the DNA sequences in or around the α_{2u} globulin gene which determine the hormonal response, by (1) mutating the α_{2u} globulin genes *in vitro*, reintroducing them into L cells and assessing the effects of the mutations on hormonal response, or (2) ligating specific fragments of the α_{2u} globulin genes to a selectable marker (for example, the HSV tk gene) in an attempt to render such genes hormonally inducible.

Similar studies to those reported here have found^{16,18} that mouse mammary tumour virus proviral DNA, cloned in λ phage and co-transformed into L cells, retains the ability to respond to dexamethasone. Thus, in at least two systems, the information necessary for hormonal response is contained in the DNA fragment used in co-transformation. Whether this indicates that the steroid hormone-receptor complex is actually interacting with the DNA (or chromatin) at a site very near or in the gene is unclear; such questions are under investigation.

I thank Doug Bishop for technical assistance, Mike Wigler and Doug Hanahan for the HSV tk gene, and Nancy Hynes and Berndt Groner for making their data available before publication. This work was supported by NIH grant AM2676B to D.T.K., and by NCI Cancer Center Research grant CA13106 to J. D. Watson.

Received 23 February; accepted 29 April 1981.

- Wigler, M. *et al. Cell* **16**, 777-785 (1979).
- Mantel, N., Boll, W. & Weissman, C. *Nature* **281**, 40 (1979).
- Lai, E. C., Woo, S. L. C., Bordelon-Riser, M. E., Fraser, T. F. & O'Malley, B. *Proc. natn. Acad. Sci. U.S.A.* **77**, 244 (1980).
- Kurtz, D. T. & Feigelson, P. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4791-4795 (1977).
- Kurtz, D. T., Chan, K.-M. & Feigelson, P. *Cell* **15**, 743-750 (1978).
- Kurtz, D. T. *J. molec. appl. Genet.* (in the press).
- Lippman, M. E. & Thompson, E. B. *J. biol. Chem.* **249**, 2483-2488 (1974).
- Hanahan, D., Lane, D., Lipsich, L., Wigler, M. & Botchan, M. *Cell* **21**, 127-139 (1980).

- Southern, E. M. *J. molec. Biol.* **18**, 503-517 (1975).
- Kurtz, D. T. & Nicodemus, C. F. *Gene* **13**, 145-152 (1981).
- Thomas, P. S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201-5205 (1980).
- Hastie, N. D., Held, W. A. & Toole, J. T. *Cell* **17**, 449 (1979).
- Kurtz, D. T., Sippel, A. E. & Feigelson, P. *Biochemistry* **15**, 1031-1036 (1976).
- Cox, R. F. *Biochemistry* **16**, 3433 (1977).
- Guyette, W. A., Matusik, R. J. & Rosen, J. R. *Cell* **17**, 1013 (1979).
- Hynes, N. E., Kennedy, N., Rahmsdorf, U. & Groner, B. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2038-2042 (1981).
- Laemmli, U. K. *Nature* **227**, 680-685 (1970).
- Buetti, E. & Diggelmann, H. *Cell* **23**, 335-346 (1981).

An aminoacyl tRNA synthetase binds to a specific DNA sequence and regulates its gene transcription

Scott D. Putney & Paul Schimmel

Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Alanine tRNA synthetase represses transcription of its own gene by binding specifically to a palindromic sequence which flanks the gene's transcription start site. Transcription repression is greatly enhanced by elevated concentrations of the cognate amino acid. The amino acid effect is caused by direct association of the ligand with the synthetase, which in turn mediates tighter binding to the DNA.

AMINOACYL tRNA synthetases carry out the first step in protein biosynthesis by attaching amino acids to the corresponding tRNAs¹. The cellular levels of these enzymes are regulated in two different ways. The first, which has been demonstrated for approximately half of the 20 synthetases, is derepression of enzyme levels during restriction of the cognate amino acid²⁻⁶. The second, termed metabolic regulation, is the dependence of synthetase levels on cellular growth rate; all synthetases studied thus far show this type of regulation⁷, as do stable RNA species^{8,9} and other proteins involved in protein synthesis.⁹⁻¹²

The work presented here deals with the molecular mechanism for the first form of regulation, although, as explained later, it may also pertain to metabolic regulation. We previously reported the cloning and partial sequencing of the gene for *Escherichia coli* Ala-tRNA synthetase¹³. This enzyme has an α_4 structure with a subunit molecular weight of 95,000 (ref. 14) making it the largest bacterial aminoacyl tRNA synthetase^{1,14}. We sequenced the 5'-coding region of the structural gene and ~250 nucleotides upstream from the start codon for the polypeptide¹³. *In vitro* transcription studies showed that initiation occurs at -79 (relative to the first base of the polypeptide start codon which is designated as +1) and that, between the transcription initiation point and the polypeptide start, there is no leader peptide-encoding region or attenuator as occurs in numerous amino acid biosynthetic operons^{13,15-22}. Thus, if transcription of *alaS* is subject to amino acid-dependent regulation, a different kind of mechanism must be operative. It was with this in mind that we expanded our initial investigations of *in vitro* transcription of this operon.

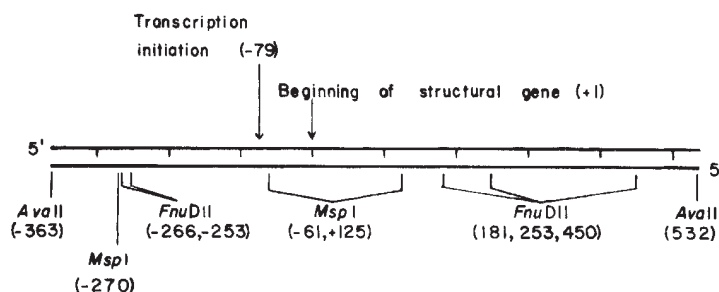


Fig. 1 Schematic diagram of the *AvaII* restriction fragment used for transcription studies. The fragment is 895 base pairs long and contains 532 nucleotides of the structural gene of *alaS*. Position +1 is assigned to the ATG at the beginning of the structural gene, and transcription is initiated with pppA at position -79. The fragment was prepared, as described previously, by purifying the plasmid pSP201 (which contains a 3.7-kilobase section of the vector pBR322 and a large segment of *alaS*) and digesting it with *AvaII*¹³. The resulting mixture was electrophoresed through a 5% (30:1) acrylamide gel and the above fragment recovered by the method of Maxam and Gilbert²³.

Ala-tRNA synthetase represses transcription of *alaS*

Figure 1 shows the 895 nucleotide-long restriction fragment used for transcription, where the two ends of the fragment are located at -363 and +532. From the sequence of the end-labelled primary transcript, it was shown previously that transcription initiates with pppA at position -79 (ref. 13). When this fragment is used as a template for transcription with [α -³²P]UTP as the sole labelled ribonucleotide, one major RNA product is synthesized (Fig. 2A, lane a). The arrow indicates the position of the transcript initiated at position -79. Figure 2A also shows the effect of 1-10 μ M concentrations of Ala-tRNA synthetase on transcription. As more synthetase is added, there is clearly a progressive decrease in the amount of primary transcript produced.

To show that Ala-tRNA synthetase does not affect promoters other than *alaS*, a 984-base pair restriction fragment containing the promoter for the histidine operon¹⁷ was transcribed in the absence and presence of the synthetase. Transcription of this fragment gives rise to an ~168-base pair leader RNA species which is heterogeneous in length at its 3'-terminus owing to stuttering of the polymerase at the string of uridines in the termination site (R. Freedman, personal communication). Figure 2B shows that, at a concentration of Ala-tRNA synthetase which represses transcription of *alaS* (lanes a, b), there is no effect on transcription from the *his* promoter (lanes c, d). Moreover, when the fragment containing the *alaS* promoter is mixed with that containing the *his* promoter, transcription proceeds cleanly from both templates (lane e) while, in the presence of the synthetase, only *alaS* transcription is repressed (lane f).

The effect on transcription of proteins other than Ala-tRNA synthetase was also investigated. Figure 2C shows the effect of 2 mg ml⁻¹ concentrations of Ala, Ile and Tyr-tRNA synthetases, as well as bovine serum albumin and lysozyme. Because all these proteins have significantly lower molecular weights than Ala-tRNA synthetase, their molar concentrations in the transcription mixture are several-fold higher than that of Ala-tRNA synthetase. Nevertheless, only the alanine-specific synthetase represses transcription of *alaS*.

Collectively, the data in Fig. 2A, B and C establish that transcription of *alaS* is specifically repressed by micromolar concentrations of the cognate synthetase.

Autoregulation of *alaS* transcription depends on alanine concentration

Figure 2D shows the effects of 0, 1 and 2 mM alanine on the Ala-tRNA synthetase-mediated repression of *alaS* transcription. Only 1 μ M of synthetase was used in the transcription mixture. As the alanine concentration is increased in the

presence of synthetase, there is a concomitant decrease in the amount of primary transcript synthesized. Many alanine concentrations were used in other experiments which utilized the synthetase at a 10-fold higher concentration and the results were analogous to those shown in Fig. 2D. Control experiments showed that there is no effect of 4 mM glycine or serine in the absence of synthetase nor of 4 mM glycine or serine in the presence of synthetase (data not shown). Although the extent of the alanine effect has not been rigorously quantified, we estimate that, in the presence of synthetase plus 1.5 mM alanine, there is more than a 20-fold decrease in the amount of transcript synthesized relative to that synthesized in the presence of synthetase alone.

To understand more fully the interaction between alanine and Ala-tRNA synthetase, we measured the K_m for alanine in the standard amino acid-dependent ATP-PP_i exchange reaction²³. In conditions of approximate pyrophosphate and ATP saturation, the K_m for alanine was found to be 1.1 mM at pH 7.5 (data not shown). (This value is 10- to 100-fold greater than the amino acid K_m s of other *E. coli* aminoacyl tRNA synthetases²⁴⁻³³.) Because the concentration of alanine required to affect tran-

scription is comparable with the independently measured K_m , we infer that the site which binds alanine for aminoacylation is the same as that which confers alanine dependence to transcription regulation.

In many of the gels shown in Fig. 2 there is a small amount of *in vitro* synthesized RNAs which migrate above the primary transcript. Previously we estimated that the major species constitutes ~90% of the RNA made from this template¹³. Transcription of these minor species is repressed in the presence of sufficient Ala-tRNA synthetase or synthetase plus alanine. This is probably because binding of the synthetase around the transcription initiation point of -79 (see below) blocks the complete synthesis of these presumably spurious, non-specific transcripts which originate upstream from -79.

Ala-tRNA synthetase binds to a region flanking transcription initiation

Because Ala-tRNA synthetase specifically affects transcription of *alaS*, it seemed plausible that the synthetase directly interacts with the DNA template itself. To test this possibility, DNA

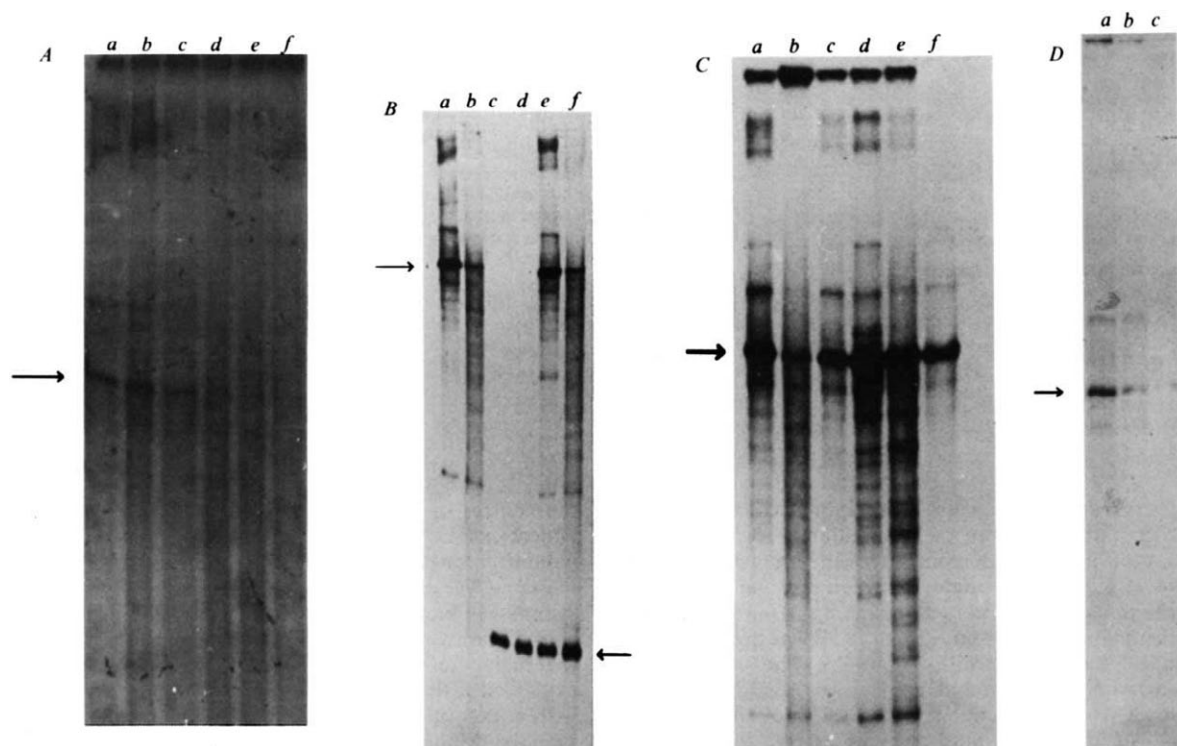


Fig. 2 A, the effect of varying concentrations of Ala-tRNA synthetase on the *in vitro* transcription of *alaS*. An autoradiograph of a denaturing polyacrylamide gel through which radioactively labelled products of transcription reactions were electrophoresed is shown. The arrow indicates the position of the ~620-nucleotide *alaS* primary transcript which begins 79 nucleotides before the start of the structural gene (see Fig. 1). The intensity of these bands indicates the amount of transcript made in each of the six reactions. The amount of Ala-tRNA synthetase present is (in μM): a, 0; b, 1.0; c, 2.5; d, 5.0; e, 7.5; f, 10.0. As the concentration of Ala-tRNA synthetase in the transcription reaction increases, the transcription of *alaS* is repressed. (The primary transcript band in lane a appears slightly less intense than that in lane b because, due to the larger width of lane a, the radioactivity covered a larger area.) Transcription reactions were performed at 37 °C in a 10- μl volume and contained 10 mM Tris-HCl pH 8.0, 10 mM MgCl₂, 0.1 mM EDTA, 150 mM KCl, 1 mM dithiothreitol, 0.5 μg purified RNA polymerase (a gift from Professor Will McClure), 50 μM each of the four ribonucleotide triphosphates (with the only labelled ribonucleotide being [α -³²P]UTP (Amersham) at a specific activity of 20 Ci mmol⁻¹), 0.1 μg of the *Ava*II fragment and the indicated amount of Ala-tRNA synthetase (purified to homogeneity using the procedure of Putney *et al.*¹⁴). The template and the synthetase were incubated for 10 min before adding the other ingredients. Reactions continued for 20 min, 40 μl of 100 $\mu\text{g ml}^{-1}$ unfractionated tRNA was then added and the mixture phenol-extracted followed by ether extraction. After ethanol precipitation the pellet was dissolved in 20 mM sodium citrate, pH 5.0, 1 mM EDTA with 9 M urea containing tracking dyes, heated to 100 °C for 1 min and electrophoresed on a 6% (20:1) polyacrylamide gel containing 7 M urea. The gel was then autoradiographed. B, the effect of Ala-tRNA synthetase on the transcription of *alaS* and on the transcription of a 984-base pair *Hind*II restriction fragment containing the promoter for the *S. typhimurium* histidine operon¹⁷. The arrow on the left indicates the position of the *alaS* primary transcript and the arrow on the right the position of the *his* leader transcript. Lanes a and b display products of transcription reactions which contained the *alaS* fragment (Fig. 1) as the template, the products in c and d are from reactions which contained the histidine operon template (a gift from Mr Richard Freedman) and the products in e and f from reactions which contained both templates. The reactions represented by lanes a, c and e had no Ala-tRNA synthetase whereas the reactions for lanes b, d and f had 10 μM Ala-tRNA synthetase. The results show that the presence of Ala-tRNA synthetase represses synthesis of the *alaS* transcript but does not affect synthesis of the *his* leader transcript. C, the effect of various proteins on the transcription of *alaS*. The *Ava*II fragment (Fig. 1) was transcribed in the presence of the following purified proteins: a, no added protein; b, 2 mg ml⁻¹ (5 μM) Ala-tRNA synthetase; c, 2 mg ml⁻¹ (18 μM) Ile-tRNA synthetase (a gift from Dr Stephen Koontz); d, 2 mg ml⁻¹ (21 μM) Tyr-tRNA synthetase (a gift from Dr Rochester Co); e, 2 mg ml⁻¹ (30 μM) bovine serum albumin (Sigma); and f, 2 mg ml⁻¹ (140 μM) lysozyme (Sigma). Only when Ala-tRNA synthetase is present during transcription is the amount of the *alaS* transcript diminished. D, the effect of alanine on the repression of *alaS* transcription by Ala-tRNA synthetase. The *Ava*II fragment (Fig. 1) was transcribed in the presence of 1 μM Ala-tRNA synthetase and various concentrations of alanine. Alanine concentrations were (in mM): a, 0; b, 1; c, 2. As the concentration of alanine in the transcription reaction increases, the repression of *alaS* transcription by Ala-tRNA synthetase is enhanced. Reactions and product analyses were performed as described in legend to A.

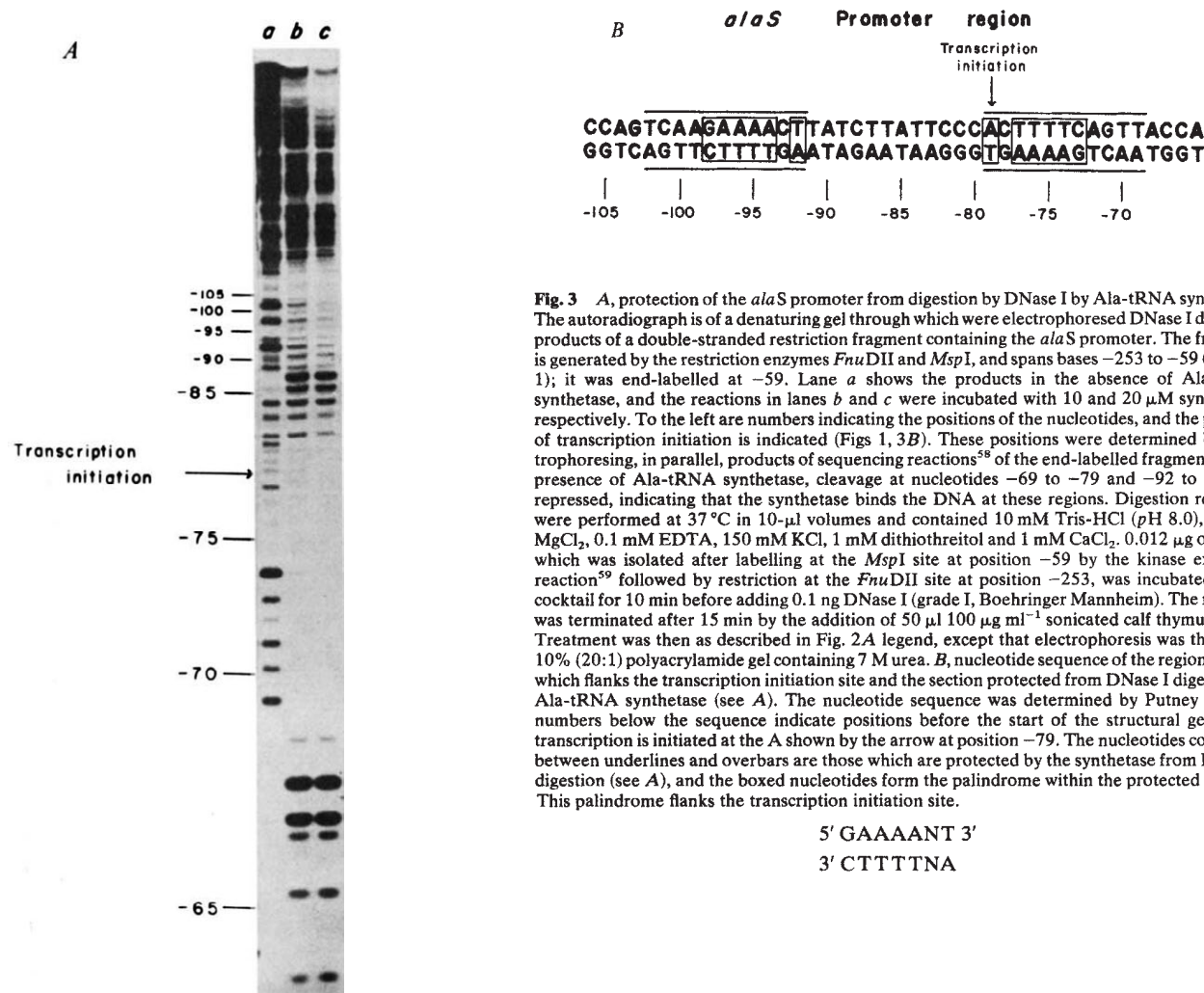


Fig. 3 A, protection of the *alaS* promoter from digestion by DNase I by Ala-tRNA synthetase. The autoradiograph is of a denaturing gel through which were electrophoresed DNase I digestion products of a double-stranded restriction fragment containing the *alaS* promoter. The fragment is generated by the restriction enzymes *FnuDII* and *MspI*, and spans bases -253 to -59 (see Fig. 1); it was end-labelled at -59. Lane *a* shows the products in the absence of Ala-tRNA synthetase, and the reactions in lanes *b* and *c* were incubated with 10 and 20 μ M synthetase, respectively. To the left are numbers indicating the positions of the nucleotides, and the position of transcription initiation is indicated (Figs 1, 3B). These positions were determined by electrophoresing, in parallel, products of sequencing reactions⁵⁸ of the end-labelled fragment. In the presence of Ala-tRNA synthetase, cleavage at nucleotides -69 to -79 and -92 to -102 is repressed, indicating that the synthetase binds the DNA at these regions. Digestion reactions were performed at 37 °C in 10- μ l volumes and contained 10 mM Tris-HCl (pH 8.0), 10 mM MgCl₂, 0.1 mM EDTA, 150 mM KCl, 1 mM dithiothreitol and 1 mM CaCl₂. 0.012 μ g of DNA, which was isolated after labelling at the *MspI* site at position -59 by the kinase exchange reaction⁵⁹ followed by restriction at the *FnuDII* site at position -253, was incubated in the cocktail for 10 min before adding 0.1 ng DNase I (grade I, Boehringer Mannheim). The reaction was terminated after 15 min by the addition of 50 μ l 100 μ g ml⁻¹ sonicated calf thymus DNA. Treatment was then as described in Fig. 2A legend, except that electrophoresis was through a 10% (20:1) polyacrylamide gel containing 7 M urea. B, nucleotide sequence of the region of *alaS* which flanks the transcription initiation site and the section protected from DNase I digestion by Ala-tRNA synthetase (see A). The nucleotide sequence was determined by Putney *et al.*¹³; numbers below the sequence indicate positions before the start of the structural gene, and transcription is initiated at the A shown by the arrow at position -79. The nucleotides contained between underlines and overbars are those which are protected by the synthetase from DNase I digestion (see A), and the boxed nucleotides form the palindrome within the protected regions. This palindrome flanks the transcription initiation site.

5' GAAAANT 3'
3' CTTTNA

footprint experiments^{34,35} were carried out with restriction fragments containing the transcription initiation region. In preliminary experiments we demonstrated an interaction of the synthetase with a section roughly between -120 and -60. To define more precisely this interaction, an *FnuDII*-*MspI* restriction fragment which runs from -253 to -59 was prepared (see Fig. 1) and 5'-end-labelled with ³²P at -59. The results of an experiment with this fragment are shown in Fig. 3A. The synthetase clearly protects nucleotides -69 through -79 from DNase I digestion; in addition, nucleotides -92 to -102 are protected. (This is more obvious on a gel on which the identical digest was electrophoresed further to improve the separation of individual fragments in the region of -85 to -110.) No other regions are as unambiguously protected by the presence of the synthetase. In addition, the synthetase enhances cleavage at certain sites, and this phenomenon has also been observed in other DNase protection studies³⁶.

Figure 3B summarizes these results. The nucleotide sequence previously determined for the region -106 to -65 is shown. Transcription initiation occurs at -79 and starting at -91 there is a sequence, TATCTTA, which fulfills the main features of a consensus sequence (Pribnow box) associated with RNA polymerase promoter sites^{13,37}. Also shown are the two nucleotide segments which bind to Ala-tRNA synthetase. These two segments flank the promoter sequence and transcription initiation point.

Within each of the protected regions is a palindromic sequence:

5' GAAAANT
CTTTNA 5'

the centres of which are 19 base pairs apart, corresponding almost exactly to two turns of a DNA B helix. Therefore, the palindromic sequences are perfectly spaced so as to interact, on the same face of the DNA, with a protein having a 2-fold axis of symmetry. Previous intersubunit cross-linking data obtained with dimethyl suberimide exclude cyclic 4-fold symmetry for the native tetrameric Ala-tRNA synthetase, but are easily explained with a 2-fold symmetry for the protein¹⁴. Moreover, with a molecular weight of 380,000, the synthetase can easily span the distance of 70-100 Å implied by the results of Fig. 3B, assuming a DNA B-helical structure. Alternatively, the palindromic sequences could arrange themselves in a hairpin-type of cloverleaf structure which interacts with the synthetase.

Implications and physiological significance

As mentioned earlier, amino acid-dependent repression of aminoacyl tRNA synthetase synthesis has been demonstrated in several cases, although no analyses have been reported for *E. coli* Ala-tRNA synthetase²⁻⁶. Based on data of Piperno and Oxender³⁸ we estimate that the intracellular concentration of alanine is between 1 and 5 mM in exponentially growing *E. coli* K12. This is substantially greater than the pool sizes of most other amino acids and, significantly, the *K_m* for alanine (for Ala-tRNA synthetase, ~1 mM in our assay conditions) is substantially higher than that reported so far for other synthetases. Thus, to a first approximation, the *K_m* and the pool size for alanine are comparable and, therefore, amino acid-dependent regulation might take place under these circumstances.

From enzyme assays of Putney *et al.*¹³, the intracellular concentration of Ala-tRNA synthetase in cells growing in rich

media is 0.4 μM ; this corresponds to 250 molecules per cell. This is in close agreement with the values of 300–900 molecules per cell for various other tRNA synthetases⁷.

At a concentration of 0.4 μM , there is little if any autogenous repression in the absence of alanine (Fig. 2A). Hence, under these circumstances, *alaS* autogenous regulation is mediated entirely by the intracellular alanine concentration. As shown in Fig. 2D, effective alanine-dependent repression of transcription occurs *in vitro* at a synthetase concentration of 1 μM and presumably at somewhat lower concentrations also. Because the transcription reaction must be carried out in the presence of ATP, we do not know whether the critical liganded species is the synthetase–amino acid, synthetase–amino acid–ATP, or synthetase–aminoacyl adenylate complex. Note, however, that the K_m for alanine was also determined in the presence of ATP (in the ATP–PP_i exchange reaction), so it can be compared directly with the concentrations of alanine required for repression of transcription. In addition, preliminary results have indicated no additional effects of tRNA^{Ala} (when added with ATP and amino acid) beyond those achieved by the amino acid alone in the transcription mixture.

The qualitative parallels between the interaction of Ala-tRNA synthetase with the *alaS* promoter region and the *lac* repressor–*lac* operator system are striking. In both cases, the protein (Ala-tRNA synthetase or *lac* repressor) binds to a palindromic sequence near the transcription initiation point^{34,39}. In both instances, a ligand (amino acid or sugar, respectively) alters the affinity of the protein for the DNA, although the effects are exerted in the opposite way (alanine enhances binding of the synthetase whereas sugar binding to the repressor weakens the interaction). In addition, both proteins are tetramers and the *lac* repressor has 2-fold symmetry⁴⁰.

The work done on the amino acid-dependent repression–derepression of aminoacyl tRNA synthetase levels^{2–6} shows that synthesis is derepressed in conditions of cognate amino acid starvation, and there are synthetase mutants which display properties consistent with amino acid-dependent autogenous regulation, such as a mutation in His-tRNA synthetase which results in an elevated K_m for histidine⁴¹. This alteration causes derepression of synthesis at histidine levels which cause repression in the wild-type strain. If histidine-dependent autoregulation is operative for this enzyme, the elevated K_m would

result in a decreased level of histidine-complexed enzyme and, consequently, less stringent repression. (There is some evidence, however, that the effector molecule is tRNA^{His} rather than free histidine^{41,42}.) Another set of examples involves mutations, located outside the synthetase structural gene, which result in increased enzyme levels. For Ser-, Val- and Leu-tRNA synthetases, these alterations are linked to the respective structural gene and are thought to be mutations in the operator locus^{43–46}. Such results could be explained by autogenous control of the type described here if the mutation occurred at the synthetase binding site in the promoter region.

Aminoacyl tRNA synthetases are also subject to metabolic regulation whereby their intracellular concentrations vary with growth rate conditions^{2–7} as do other species involved in protein synthesis^{8–12}. Neidhardt *et al.* have proposed a mechanism for both amino acid-dependent repression and metabolic control involving just the substrates and products of the aminoacylation reaction². This mechanism would also explain amino acid-dependent repression of aminoacyl tRNA synthetase synthesis. On the other hand, there is evidence that guanosine tetraphosphate is related to aminoacyl tRNA synthetase regulation^{47,48}, as well as to aminoacyl tRNA synthesis^{49,50} and enzyme synthesis from operons which are subject to amino acid control via the attenuator mechanism^{51,52}. At present our data only pertain to the mechanism of amino acid-dependent repression; however, with suitable modification, the Neidhardt *et al.* proposal might be harmonized with our data to explain both amino acid-dependent repression and metabolic control, but it is likely that more considerations are involved than have been explored here.

Several ribosomal proteins, the levels of which are under metabolic regulation, are also autogenously controlled^{51–57}. However, in these cases, much of the effect seems to be at the translational level (although transcription may also be affected⁵⁷), and no co-repressor has been found as for Ala-tRNA synthetase. We have not investigated the possibility of autoregulation of translation of Ala-tRNA synthetase, but in view of the effective transcriptional control which operates we question the need for (but do not rule out) an additional translation based regulatory mechanism.

This work was supported by NIH grant GM15539. S.P. was a NIH Predoctoral Trainee from 1977 to 1980.

Received 29 December 1980; accepted 30 March 1981.

- Schimmel, P. R. & Söll, D. A. *Rev. Biochem.* **48**, 601–648 (1979).
- Neidhardt, F. C., Parker, J. & McKeever, W. G. A. *Rev. Microbiol.* **29**, 215–250 (1975).
- Brenchley, J. E. & Williams, L. S. A. *Rev. Microbiol.* **29**, 251–274 (1975).
- La Rossa, R. & Söll, D. in *Transfer RNA* (ed. Altman, S.) 136–167 (MIT Press, Cambridge, Massachusetts, 1978).
- Umbarger, H. E. in *Transfer RNA: Biological Aspects* (eds Söll, D., Abelson, J. & Schimmel, P.) 453–467 (Cold Spring Harbor Laboratory, New York, 1980).
- Morgan, S. D. & Söll, D. *Prog. Nucleic Acid Res. molec. Biol.* **21**, 181–207 (1978).
- Neidhardt, F. C., Bloch, P. L., Pedersen, S. & Reeh, S. J. *Bact.* **129**, 378–387 (1977).
- Skjold, A. C., Jauarez, H. & Hedgcock, C. J. *Bact.* **115**, 177–187 (1973).
- Kjeldgaard, N. O. & Gausing, K. in *Ribosomes* (eds Nomura, M., Tissieres, A. & Lengyel, P.) 369–392 (Cold Spring Harbor Laboratory, New York, 1974).
- Gordon, J. *Biochemistry* **9**, 912–917 (1970).
- Pedersen, S., Bloch, P., Reeh, S. & Neidhardt, F. *Cell* **14**, 179–190 (1978).
- Herendeen, S., Van Bogelen, R. & Neidhardt, F. J. *Bact.* **139**, 185–194 (1979).
- Putney, S. D., Melendez, D. & Schimmel, P. R. J. *biol. Chem.* **256**, 205–211 (1981).
- Putney, S. D., Sauer, R. T. & Schimmel, P. R. J. *biol. Chem.* **256**, 198–204 (1981).
- Bennett, G. N., Schweingruber, M. E., Brown, K. D., Squires, C. & Yanofsky, C. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2351–2355 (1976).
- Di Nocera, P. P., Blasi, F., Di Lauro, R., Frunzio, R. & Bruni, C. B. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4276–4280 (1978).
- Barnes, W. M. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4281–4285 (1978).
- Zurawski, G., Brown, K., Killingley, D. & Yanofsky, C. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4271–4275 (1978).
- Gardner, J. F. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1706–1710 (1979).
- Gemmill, R. M., Wessler, S. R., Keller, E. B. & Calvo, J. M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4941–4945 (1979).
- Nargang, F. E., Subrahmanyam, C. S. & Umbarger, H. E. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1823–1827 (1980).
- Lawther, R. P. & Hatfield, W. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1862–1866 (1980).
- Calendar, R. & Berg, P. in *Procedures in Nucleic Acids Research* (eds Cantoni, G. L. & Davies, D. R.) 384–399 (Harper & Row, New York, 1966).
- Calendar, R. & Berg, P. *Biochemistry* **5**, 1690–1695 (1966).
- Kalousek, F. & Konigsberg, W. *Biochemistry* **13**, 999–1006 (1974).
- Lemoine, F., Waller, J. P. & Rapenbusch, R. *Eur. J. Biochem.* **4**, 213–221 (1968).
- Katze, J. R. & Konigsberg, W. J. *biol. Chem.* **245**, 923–930 (1970).
- Joseph, D. & Meunch, K. J. *biol. Chem.* **246**, 7602–7609 (1971).
- Hirshfield, I. & Bloemers, H. P. J. *J. biol. Chem.* **244**, 2911–2916 (1969).
- Rouget, P. & Chapeville, F. *Eur. J. Biochem.* **14**, 498–508 (1970).
- Yaniv, M. & Gros, F. J. *molec. Biol.* **44**, 1–15 (1969).
- Cole, F. X. & Schimmel, P. R. *Biochemistry* **9**, 480–489 (1970).
- Ostrem, D. & Berg, P. *Biochemistry* **13**, 1338–1348 (1974).
- Galas, D. J. & Schmitz, A. *Nucleic Acids. Res.* **5**, 3157–3170 (1978).
- Johnson, A. D., Meyer, B. J. & Ptashne, M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5061–5065 (1979).
- Schmitz, A. & Galas, D. J. *Nucleic Acids Res.* **6**, 111–137 (1979).
- Rosenberg, M. & Court, D. A. *Rev. Genet.* **13**, 319–353 (1979).
- Piperno, J. R. & Oxender, D. L. J. *biol. Chem.* **243**, 5914–5920 (1968).
- Gilbert, W. & Maxam, A. *Proc. natn. Acad. Sci. U.S.A.* **70**, 3581–3584 (1973).
- Steitz, T. A., Richmond, T. J., Wise, D. & Engelman, D. *Proc. natn. Acad. Sci. U.S.A.* **71**, 593–597 (1974).
- McGinnis, E. & Williams, L. J. *Bact.* **111**, 739–744 (1972).
- McGinnis, E. & Williams, L. J. *Bact.* **109**, 505–511 (1972).
- Pizer, L. I., McKittrick, J. & Tosa, T. *Biochem. biophys. Res. Commun.* **49**, 1351–1357 (1972).
- Clarke, S. J., Low, B. & Konigsberg, W. J. *Bact.* **113**, 1096–1103 (1973).
- Baer, M., Low, K. B. & Söll, D. J. *Bact.* **139**, 165–175 (1979).
- La Rossa, R., Vogeli, G., Low, K. B. & Söll, D. J. *molec. Biol.* **117**, 1033–1048 (1977).
- Hirshfield, I. N., Liu, C. & Yeh, F. J. *Bact.* **131**, 589–597 (1977).
- Wirth, R., Buckel, P. & Böck, A. *FEBS Lett.* **83**, 103–106 (1977).
- Cashel, M. & Gallant, J. in *Ribosomes* (eds Nomura, M., Tissieres, A. & Lengyel, P.) 733–745 (Cold Spring Harbor Laboratory, New York, 1974).
- Block, R. & Haseltine, W. A. in *Ribosomes* (eds Nomura, M., Tissieres, A. & Lengyel, P.) 747–761 (Cold Spring Harbor Laboratory, New York, 1974).
- Morse, D. E. & Morse, A. N. J. *molec. Biol.* **103**, 209–226 (1976).
- Stephens, J. C., Artz, S. W. & Ames, B. N. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4389–4393 (1975).
- Yates, J. L., Arfsten, A. E. & Nomura, M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1837–1841 (1980).
- Fill, N. P., Friesen, J. D., Downing, W. L. & Dennis, P. P. *Cell* **19**, 837–844 (1980).
- Dean, D. & Nomura, M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3590–3594 (1980).
- Yates, J. L. & Nomura, M. *Cell* **21**, 517–522 (1980).
- Zengel, J. M., Mueckl, D. & Lindahl, L. *Cell* **21**, 523–535 (1980).
- Maxam, A. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560–564 (1977).
- Berkner, K. L. & Folk, W. R. J. *biol. Chem.* **252**, 3176–3189 (1977).

LETTERS TO NATURE

Newtonian gravity measurements impose constraints on unification theories

G. W. Gibbons & B. F. Whiting

Department of Applied Mathematics and Theoretical Physics,
University of Cambridge, Silver Street, Cambridge CB3 9EW, UK

Theories which attempt to unify gravity with the other forces of nature can be coarsely classified according to the mass scale of the new particles they introduce or equivalently the length scale at which new phenomena occur. This mass scale can be expressed as $m_p(m_H/m_p)^n$, where m_H is a typical hadron mass and m_p is the Planck mass. In most current theories n is 0 or 1. However, in some theories $n = 2$, which offers the possibility of experimental consequences at kilometre scales. Here using satellite and geophysical data we place constraints on such theories and find that they are not viable unless $m_H > 10^3$ GeV.

These theories¹⁻⁷ predict deviations from Newton's inverse square law at length scales $L \approx (m_p/m_H)\hbar/m_Hc$ where m_p is the Planck mass

$$((c\hbar/G)^{1/2} \approx 2.2 \times 10^{-5} \text{ g} = 1.2 \times 10^{19} \text{ GeV})$$

An L of $\approx (1 \text{ GeV}/m_H)^2 \text{ km}$ is the Compton wavelength of a particle of mass

$$\mu = (m_H/1 \text{ GeV})^2 10^{-10} \text{ eV} \approx (m_H/1 \text{ GeV})^2 10^{-6} \text{ K}$$

If m_H is the neutron mass we obtain the order of magnitude for the radius of a neutron star. In the theories of Fujii¹⁻⁴ where the masses of elementary particles and the dimensional nature of the gravitational coupling constant G have their common origin in a spontaneous breakdown of dilation invariance, μ is the mass of a scalar particle which couples to ordinary matter with gravitational strength. Zee^{5,6} derives Newton's constant in a similar way, but the mass of the resultant scalar particle in his theory would seem to be much heavier than the masses envisaged here. Scalars with low mass can also be a consequence of the super-Higgs effect in supergravity theories. In extended supergravity theories the further possibility arises of vector particles with a mass $\approx \mu$. These would also couple to ordinary matter with gravitational strength but would give rise to a repulsive force⁹⁻¹². Note that Majorana Fermion mass terms of this magnitude have been suggested for neutrinos arising from a gravitationally-induced four-Fermi interaction¹³ and for the spin 3/2 gravitino particle in supergravity theories through a supersymmetric Higgs Mechanism¹⁴. Because of supersymmetry one might expect bosonic partners with mass of order μ . Such a spin 1 partner for the gravitino has been suggested by Fayet¹⁵.

Length scales of 1 km lie between the distance range of laboratory measurements of Newton's constant, G , and astronomical distances over which Newton's law has most accurately been checked. It has been argued¹⁶ that our experimental knowledge of gravitational forces between 1 m and 10 km is so poor that it allows a considerable difference between the laboratory measured gravitational constant and its value on astronomical scales—an effect predicted in theories of the type alluded to above. Even at laboratory scales Long¹⁷ has reported non-newtonian behaviour. The null experiment of Spero *et al.*⁸ conflicts with Long's result unless some mechanism like that suggested by Long¹⁹ operates. However, conventional vacuum polarization effects in quantum gravity do not lead to the behaviour required by Long¹⁹; such effects are insignificant²⁰. We point out that existing data from satellite and

geodesy measurements^{21,22} together with the results of a recent mine experiment²³ and those in refs 18, 24 place stringent limits on the parameters appearing in theories of the Fujii or Scherk types.

If gravitation is due to the exchange of particles of Compton wavelength λ_i , and coupling α_i , the potential energy of two masses m and m' at a separation r is V where

$$V = -G_\infty \frac{mm'}{r} \left(1 + \sum_{i=1}^N \alpha_i \exp(-r/\lambda_i) \right) \quad (1)$$

As the additional Yukawa terms are negligible for $r \gg \max \lambda_i$, G_∞ is the gravitational constant at very large distances. For $r \ll \min \lambda_i$ the newtonian law also applies with a gravitational constant

$$G_0 = G_\infty \left(1 + \sum_{i=1}^N \alpha_i \right)$$

Negative values of α_i correspond to repulsions caused by the exchange of vector particles and positive α_i to attractions due to the exchange of scalar particles. Potentials of the form of equation (1) occur in several scalar-tensor theories of gravity^{25,26} in which the mass of the scalar particle remains unspecified. Such potentials also arise in higher derivative theories of gravity²⁷⁻²⁹, the mass usually being thought of as having a planckian or possibly hadronic value. According to De Witt²⁸ the value is to be determined experimentally. Stelle²⁹ has shown that theories predict $\alpha = \frac{1}{3}$ for the massive scalar and $\alpha = -\frac{4}{3}$ for the massive tensor ghost particle. Fujii's first theory¹ also predicted $\alpha = \frac{1}{3}$.

For the present purposes we consider $N = 1$, so that

$$V = -G_\infty \frac{mm'}{r} (1 + \alpha \exp(-r/\lambda)) \quad (2)$$

The effective gravitational constant $G(r)$ is given by

$$G(r) = G_\infty (1 + \alpha (1 + r/\lambda) \exp(-r/\lambda)) \quad (3)$$

A measured value of the ratio $G(r_1)/G(r_2) = \beta$ at the distances r_1 and r_2 constrains α and λ to lie on a curve in (α, λ) space given by

$$\alpha = \frac{(\beta - 1)}{[(1 + r_2/\lambda) \exp(-r_2/\lambda) - \beta(1 + r_1/\lambda) \exp(-r_1/\lambda)]} \quad (4)$$

Varying β within the limits of experimental errors causes the curve to sweep out an allowed region in the (α, λ) plane. Spero *et al.*¹⁸ used this procedure to obtain the region allowed by their results and by those of Long¹⁷; these are plotted in Fig. 1 for positive and negative α . Figure 1 also shows the regions allowed by the experiment of Panov and Frontov²⁴, which effectively restrict $|\alpha|$ to be $< 10^{-2}$ out to 3 m. (The results of Yu *et al.*³⁰ and Hirakawa *et al.*³¹ provide poorer limits and are not included.)

These experiments constrain $G(r)$ at laboratory scales. Mikkelsen and Newman have provided limits from lunar surface gravity and Mercury and Venus flyby data which constrain $G(r)$ at the upper end of the range we are considering. These are also shown in Fig. 1 from which it can be seen that for $3 \text{ m} < \lambda < 10^3 \text{ km}$ α is very poorly constrained. To fill this gap we turn to the data available from satellites and geodesy^{21,22}, using essentially independent computations of equatorial gravity from an extensive net of gravity anomaly measurements and from satellite data for the Earth's potential. Mikkelsen and Newman regarded data of this sort as unreliable; they instead considered lunar surface gravity data which, because of the smaller size of the Moon, gives a much tighter constraint on α for a given error of measurement. We believe that more recent data avoids the objections raised for the Earth in ref. 16.

The potential at a distance r from the centre of a uniform sphere of density ρ and radius R due to the Yukawa contribu-

tion to gravity is

$$\frac{4\pi}{3} \cdot \frac{G_{\infty} \rho R^3}{r} \alpha f(R/\lambda) \exp\{(R-r)/\lambda\} \quad (5)$$

where

$$f(x) = 3e^{-x}(x \cosh(x) - \sinh(x))x^{-3} \quad (6)$$

is a 'form factor' which takes into account the fact that only a region of the order of λ from the field point contributes to the Yukawa part of gravity. For a non-uniform sphere ρ should be replaced by a suitably averaged density $\bar{\rho}$ within a few λ of the field point. The acceleration due to gravity at a height h above the surface of the Earth is therefore

$$g(h) = G_{\infty} M_{\oplus} (R+h)^{-2} \times [1 + \gamma \alpha f(R/\lambda) (1 + (R+h)/\lambda) \exp(-h/\lambda)] \quad (7)$$

where $M_{\oplus} = 4\pi \rho R^3/3$ is the mass of the Earth (ρ now being the mean density) and $\gamma = \bar{\rho}/\rho$. On the surface ($h=0$)

$$g(h) \rightarrow g(0) = G_{\infty} M R^{-2} [1 + \gamma \alpha f(R/\lambda) (1 + R/\lambda)] \quad (8)$$

At a height $h \gg \lambda$ we have

$$g(h) \rightarrow g_{\infty}(r) = G_{\infty} M (R+h)^{-2} \quad (9)$$

Thus the extrapolation from measurements of gravity at large heights above the Earth to the surface will produce a fractional error

$$\varepsilon = (g_{\text{measured}} - g_{\text{extrapolated}}) / g_{\text{extrapolated}} \quad (10)$$

$$= \gamma \alpha f(R/\lambda) (1 + R/\lambda) \quad (11)$$

If $R \gg \lambda$ this gives

$$\varepsilon = 3\alpha\lambda\gamma/2R \quad (12)$$

The accuracy with which terrestrial and satellite measurements of gravity agree has been discussed by Rapp^{21,22} who provides estimates of normal gravity at the Equator, γ_e , derived from satellite models, ranging from 978,031.3 to 978,033.0 mGal (1 Gal = 1 cm s⁻²) and also determines γ_e , from an extensive net of 1° × 1 mean free air anomalies, as 978,030.9 mGal with an estimated standard error of ±1 mGal. Thus satellite and terrestrial values agree to within ~2 p.p.m. Rapp's adjusted value of γ_e , 978,031.69 mGal is based on a best set of satellite data and includes his geodetic determination of γ_e . His computation (R. H. Rapp, personal communication) enables the dependence of this adjusted value on the geodetic data to be removed and the resultant estimate, $\gamma_e = 978,032.085$ mGal agrees with the geodetic data to within 1.2 mGal. All Rapp's calculations have removed the effect of the atmosphere. Thus:

$$\varepsilon = (978,030.9 - 978,032.085) / 978,032.085 = -1.2 \times 10^{-6}$$

Rapp has suggested (personal communication) that current estimates of the agreement may be a factor of 10 better than this.

Mikkelsen and Newman's¹⁶ objected to geophysical data because the local agreement between measured and calculated gravity anomalies may be a factor of 10³ worse than the agreement of these globally averaged quantities. This is too pessimistic; Rapp²² shows that the agreement is within 10 mGal.

Rapp computes the potential coefficients in a multipole expansion of the Earth's gravitational field from 5° × 5° mean gravity anomalies and compares his results with the coefficients of the Earth model GEM 7 which were obtained using satellites. The mean square difference in gravity anomalies, δg , calculated allowing for various terrain corrections, depended on the number of harmonics used. For a maximum degree of 3, δg did not exceed 2.2 mGal; as the maximum degree increased these differences increased, but even for a maximum degree of 16 δg never exceeded 7.3 mGal. These relatively small discrepancies for local anomalies indicate that an agreement of as little as 10⁻⁶ is reasonable for a globally averaged quantity and a discrepancy larger than 10⁻⁵ is unacceptable. The constraints obtained in both these cases are shown in Fig. 1. The larger value of the Earth's radius is more than compensated for by the greater

precision of the measurements compared with the lunar surface gravity measurements. The constraint, $\alpha < 10^{-2}$, applies if $\lambda > 1$ km for $\varepsilon = 10^{-6}$, and if $\lambda > 10$ km for $\varepsilon = 10^{-5}$.

We now discuss the effect of the Yukawa addition to gravity on the variation of gravity with depth below the surface of the Earth. This variation may be measured by experiments in mines²³ or in submarines³². The acceleration due to gravity at a depth d below the surface of a homogeneous sphere of radius R and density ρ is

$$g(R-d) = \alpha \frac{4\pi}{3} G_{\infty} \rho (R-d) \times \times (1 + R/\lambda) \exp(-d/\lambda) f((R-d)/\lambda) \quad (13)$$

Assuming λ/R and d/R are both small we obtain

$$g(R-d) - g(R) = 2\pi G_{\infty} \rho \alpha \lambda (\exp(-d/\lambda) - 1) \quad (14)$$

When applying equation (14) to an inhomogeneous Earth the appropriate density ρ is ρ_m , obtained as some average of the density of the material through which the descent is made, including material to a distance λ below the depth d .

The standard newtonian change in gravity is given by

$$g_N(R-d) - g_N(R) = -\frac{4\pi}{3} G_{\infty} (3\rho_m - 2\rho_{\oplus}) d \quad (15)$$

where $\rho_{\oplus}$ is the mean density of the Earth.

Equation (15) is accurate to an order of (d/R) . To this order of accuracy equation (5) or (13) both give negligible contribution and give, for gravity at the surface of the Earth,

$$g_0 = g(R) = \frac{4\pi}{3} G_{\infty} \rho_0 R \quad (16)$$

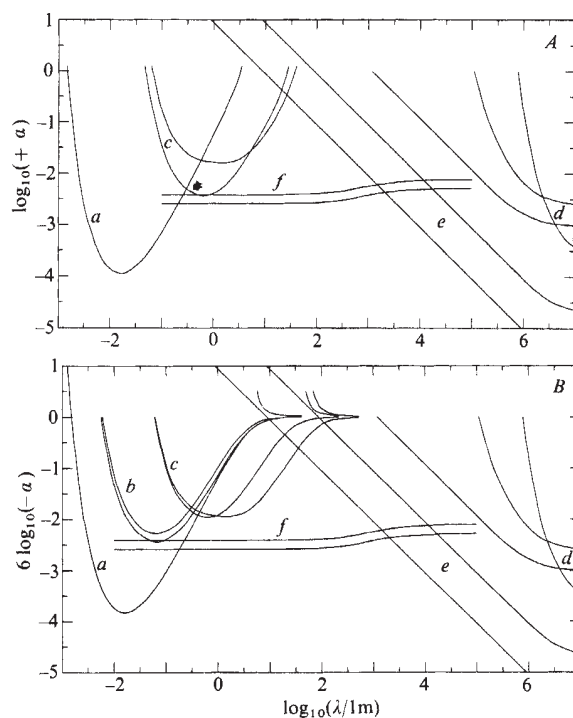


Fig. 1 A, $\log_{10}(+\alpha)$ plotted against $\log_{10}(\lambda/1\text{m})$ for $\alpha > 0$. B, $\log_{10}(-\alpha)$ plotted against $\log_{10}(\lambda/1\text{m})$ for $\alpha < 0$. Curves: a, from data and analysis in ref. 18 but calculated at the level of 1 s.d.; b, from data in ref. 17, α must lie between the curves; c, from data in ref. 24 for two separate ranges of experiment; d, taken from ref. 16, giving curves obtained from results for lunar surface gravity, and Mercury and Venus fly-bys; e, resulting from comparison of satellite and geodesy data²¹; the upper curve is assuming there is agreement to 0.1 p.p.m., the lower 1 p.p.m.; f, calculated for a mine experiment to 1% accuracy (upper curve) and a submarine experiment to 0.1% accuracy, both to a depth of 1 km. Neither curve can be extended indefinitely for small or large λ . Note that α is restricted to lie below each curve.

Expressing results in terms of laboratory measured values of Newton's constant, $G_0 = (1 + \alpha)G_\infty$, and surface gravity, g_0 , the total change in gravity is

$$\Delta g(\alpha, d) = -4\pi d \left(\frac{\rho_M G_0}{1 + \alpha} - \frac{g_0}{2\pi R} \right) - \frac{\alpha}{1 + \alpha} 2\pi G_0 \rho_M \lambda (1 - \exp(-d/\lambda)) \quad (17)$$

If ε is the fractional difference between this and the newtonian value,

$$\varepsilon = (\Delta g(\alpha, d) - \Delta g(0, d)) / \Delta g(0, d) \quad (18)$$

An experiment giving a value for ε again defines a curve in the (α, λ) plane according to

$$\alpha = -\varepsilon \left(1 - \frac{g_0}{2\pi G_0 \rho_M R} \right) / \left(\varepsilon \left(1 - \frac{g_\infty}{2\pi G_\infty \rho_M R} \right) + 1 - \frac{\lambda}{2d} (1 - \exp(-d/\lambda)) \right) \quad (19)$$

Experimental limits on ε then confine the theory to a restricted region of the (α, λ) plane. Figure 1 shows the curves obtained by setting $\varepsilon = 0.01$, $\rho = 2.65$ as a representational value for crustal rock and $\varepsilon = 0.001$, $\rho = 1.03$ for seawater, assuming an experiment carried out to a depth of 1 km. Mine experiments accurate to only a few per cent can give very useful limits to α in the range of λ where α is otherwise poorly constrained. Because of factors related to the density differences involved a submarine experiment to the same depth as a mine experiment will need ~ 10 times greater accuracy before it can put a better constraint on α . Stacey has reported results from an experiment to a depth of 950 m in a mine at Mount Isa, Queensland giving $|\varepsilon| < 0.05$.

We conclude that there is very little scope for a theory which allows deviations of $>1\%$ from Newton's law of gravitational attraction on laboratory or larger length scales. Good mine or submarine experiments will further constrain these theories. Further large scale experiments are essential to improve bounds on α between 1 m and 10 km.

We thank A. H. Cook, D. Gubbins, S. W. Hawking, and R. Newman for useful discussions and F. Stacey, P. Bender and R. H. Rapp for useful correspondence.

Received 3 November 1980; accepted 24 April 1981.

1. Fujii, Y. *Nature phys. Sci.* **234**, 5-7 (1971).
2. Fujii, Y. *Ann. Phys.* **69**, 494-521 (1972).
3. Fujii, Y. *Phys. Rev. D* **9**, 874-876 (1974).
4. Fujii, Y. *Gen. Relat. Gravitat.* **6**, 29-34 (1975).
5. Zee, A. *Phys. Rev. Lett.* **42**, 417-521 (1979).
6. Zee, A. *Phys. Rev. Lett.* **44**, 703-706 (1979).
7. Ellis, J., Galliard, M. K. & Zumino, B. *Phys. Lett.* **94B**, 343-348 (1980).
8. Fujii, Y. & Nishino, N. *Z. Phys. C* **2**, 247-252 (1979).
9. Scherk, J. *Phys. Lett.* **88B**, 265-267 (1979).
10. Scherk, J. in *Supergravity* (eds Freedman, D. & van Nieuwenhuizen, P.) 43-51 (North-Holland, Amsterdam, 1979).
11. Scherk, J. in *Mathematical Problems in Theoretical Physics* (ed. Osterwalder, K.) (Springer Lecture Notes in Physics No. 116, 1980).
12. Scherk, J. in *Unification of the Fundamental Particle Interactions* (eds Ferrara, S., Ellis, J. & van Nieuwenhuizen, P.) 381-410 (Plenum, New York, 1981).
13. Barber, R., Ellis, J. & Galliard, M. K. *Phys. Lett.* **90B**, 253-257 (1980).
14. Deser, S. & Zumino, B. *Phys. Rev. Lett.* **38**, 1433-1436 (1977).
15. Fayet, P. *Phys. Lett.* **95B**, 285-289 (1980).
16. Mikkelsen, D. R. & Newman, N. J. *Phys. Rev. D* **16**, 919-926 (1977).
17. Long, D. R. *Nature* **260**, 417-418 (1976).
18. Spero, R. *et al. Phys. Rev. Lett.* **44**, 1645-1648 (1980).
19. Long, D. R. *Nuovo Cim.* **85B**, 252-256 (1980).
20. Duff, M. J. *Phys. Rev. D* **9**, 1837-1839 (1974).
21. Rapp, R. H. *Geophys. Res. Lett.* **1**, 35-38 (1974).
22. Rapp, R. H. *Bull. Geod.* **51**, 301-323 (1977).
23. Stacey, F. D. *et al. Phys. Rev. D* **23**, 1683-1692 (1981).
24. Panov, V. I. & Frontov, V. N. *Soviet Phys. JETP* **50**, 852-856 (1979).
25. O'Hanlon, J. *Phys. Rev. Lett.* **29**, 137-138 (1972).
26. Acharya, R. & Hogan, P. A. *Lett. Nuovo Cimento* **6**, 668-672 (1973).
27. Pechlana, E. & Sexl, R. *Commun. Math. Phys.* **2**, 165-175 (1966).
28. de Witt, B. *Phys. Rep.* **19C**, 295-357 (1975).
29. Stelle, K. *Gen. Relat. Gravitat.* **9**, 353-371 (1978).
30. Yu, H. T. *et al. Phys. Rev. D* **20**, 1813-1815 (1979).
31. Hirakawa, H. *et al. Nature* **283**, 184-185 (1980).
32. Stacey, F. *Geophys. Res. Lett.* **5**, 377-378 (1978).

Meteoric compounds in the stratosphere

A. C. Aikin

Laboratory for Planetary Atmospheres, Goddard Space Flight Center, Greenbelt, Maryland 20771, USA

The presence of gaseous metal oxides and hydroxides originating from incoming meteoric material has been suggested to explain stratospheric ion composition¹ and as chemical sinks for chlorine species². Recent ion composition data can be explained without recourse to metal hydroxides³⁻⁵, although the results do not preclude gaseous metal compound concentrations of 10^4 cm^{-3} at 35 km. However, theoretical considerations suggest that gaseous meteoric material coagulates to form particulate matter⁶. No direct measurements of metal compounds exist. Photoionization of meteoric compounds such as NaOH and CaOH in the 40-70-km altitude region produces a diurnally varying source of ionization in addition to nearly constant cosmic ray ionization. Ion pair production can be several times that due to cosmic radiation depending on sunspot cycle, altitude, time of day and meteor activity. It is suggested here that the behaviour of the ion density should be studied to determine the existence of gaseous metal oxides. The available ion density and conductivity data are insufficient to demonstrate conclusively the existence of such compounds, although unexplained increases are observed. The presence of meteoric compounds, would imply an electrical conductivity which is related to both solar and meteor activity and would thus be an important factor in the physical processes involved in those Sun-weather relationships which depend on atmospheric electrical conductivity.

Table 1 lists several of the more abundant elements contained in meteoric material. The abundance is given relative to silicon, which is taken to be 100. Also listed are compounds, which are likely to be formed as the meteoric element passes through the atmosphere. The wavelength equivalent of the ionization threshold of the compounds is also given. Except for NaOH these values are based on the compilation by Rosenstock *et al.*⁷. The ionization potential for NaOH is estimated by adding H to NaO for which the ionization threshold is known: this addition should not raise the ionization threshold. Candidates for ionization below 70 km include CaO, CaOH, NaO and NaOH. In the case of Na the hydroxyl, NaOH, is predicted to be the principal compound below 80 km chiefly because of the reactions

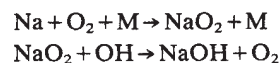


Table 1 Elemental abundance and ionization threshold of meteoric compounds

Element	Abundance	Compound	Ionization threshold (Å)
Si	100	SiO	1,062.5
Fe	74	FeO	1,550.0
Ni	3.9	NiO	1,305.3
Cr	1.1	CrO	1,476.2
Al	5.9	AlO	1,305.3
Na	4.3	NaO	1,907.7
		NaOH	1,907.7
K	0.3	KOH	1,653.3
Ca	5.5	CaO	1,907.7
		CaOH	2,101.7

All values except for NaOH are based on the compilation in ref. 7.

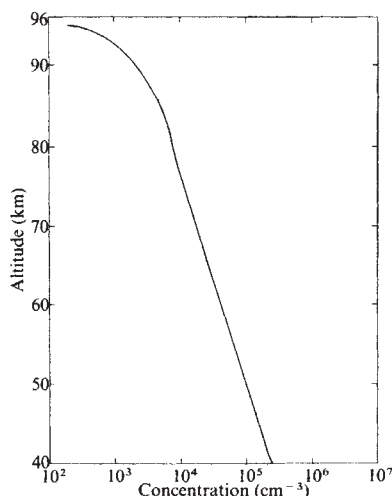


Fig. 1 Adopted altitude distribution of NaOH or CaOH.

Calculations for atomic sodium diffusing downward from its observed peak at 90 km, while reacting photochemically but with no loss due to particulate formation, are in general agreement with similar calculations⁸: NaOH is the principal sodium compound.

Because calcium can have nearly the same meteoric abundance as sodium⁹, the total density distribution of calcium compounds will follow that of sodium (neutral chemistry for calcium has not been established). If a reaction sequence analogous to that for sodium occurs, then one expects CaOH to be the principal component below 70 km. For calculations of the ionization rate the distribution shown in Fig. 1 will be adopted and be assumed to apply to either CaOH or NaOH.

The photoionization production rate for a constituent of concentration $[M]$ is derived using the expression

$$q = 10^{-18} \sum_{\lambda > \lambda_0} [M] Q_{\lambda} e^{-\tau_{\lambda}}$$

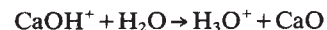
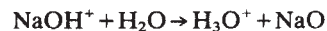
where Q_{λ} is the solar flux at wavelengths less than the ionization threshold, λ_0 (1,908 Å for NaOH, 2,100 Å for CaOH). The optical depth τ_{λ} is the product of column densities and absorption cross-sections. Ozone and molecular oxygen are the absorbers. At 2,000 Å there is minimum in the O_3 cross-section which allows the penetration of this radiation to photoionize CaOH. The present calculation is for a solar zenith angle of 30°. Photoionization cross-section is assumed to be 10^{-18} cm^2 . The total ion-pair production function shown in Fig. 2 has diurnal variation superimposed on the nearly constant cosmic-ray ionization rate. Ionization by cosmic radiation varies by a factor of two with solar cycle and an order of magnitude with geomagnetic latitude^{10,11}. Both the minimum amount of cosmic-ray ionization, geomagnetic equator at sunspot maximum, and nearly the maximum amount, high latitude sunspot minimum, are given in Fig. 2. Any photoionization of meteoric natural adds to the cosmic radiation. Solar $\text{Ly}\alpha$ ionization of nitric oxide is the ionizing agent above 70 km.

For circumstances at low geomagnetic latitudes CaOH^+ and NaOH^+ can be dominant ions formed first by ionization. At high geomagnetic latitudes the influence of photoionization will be minimized due to the combination of lower average solar zenith angles and larger cosmic-ray ionization.

The hypothesis that meteoric material is adding to the ionization level of the atmosphere below 70 km is consistent with observations. Widdel *et al.*¹² have observed an increase in ion density above 40 km with a maximum of 10^4 cm^{-3} at 50 km. The loss rate in this altitude region is due to ion-ion recombination between positive and negative ions. Smith *et al.*¹³ have measured

the ion-ion recombination rate for negative ions observed in the atmosphere and water cluster ions. A mean value of $5 \times 10^{-8} \text{ cm}^3 \text{ s}^{-1}$ is observed. No measurements are available for ion-ion recombination involving meteor material-based ions, but a value in the $5 \times 10^{-8} \text{ cm}^3 \text{ s}^{-1}$ range cannot be excluded. The observed ion density cannot be explained by cosmic radiation alone using this value of the ion-ion recombination coefficient. The ion pair production rate will also vary in direct proportion to the meteor flux.

Ion composition measurements show that water cluster ions predominate to 55 km with other ions, increasing in importance at lower altitudes so that by 40 km 50% or more of the ions are not $\text{H}^+(\text{H}_2\text{O})_n$. If CaOH^+ or NaOH^+ are formed in large quantities then loss mechanisms must transfer these ions to match the observed composition. Possible reactions include



The ion composition measurements above 50 km were conducted at a geomagnetic latitude of 69° often at high solar zenith angles. These are precisely the conditions where the contribution by photoionized meteoric material is expected to be minimal so that more H_3O^+ is expected.

An additional constraint is that the densities of ionizable constituents such as NaO, CaO, CaOH and NaOH must be low enough in the D region so that $\text{Ly}\alpha$ ionization remains the primary mechanism for ion formation between 70 and 85 km. Diurnal studies of electron density support $\text{Ly}\alpha$. Sunrise variations are more difficult to interpret due to photodetachment of negative ions and lack of precise data. Long-wave UV radiation will penetrate to 70 km before $\text{Ly}\alpha$ and produce ionization at an earlier time. The data of Mitchell *et al.*¹⁴ indicate anomalous increases in conductivity at sunrise. Data of Croskey *et al.*¹⁵ show an increase in equatorial conductivity above 60 km compared with mid-latitude values. Both these measurements are consistent with an additional source of ionization.

There is, therefore, no direct measurement of NaOH or other metal compounds postulated by Murad *et al.*² to react chemically with stratospheric chlorine. Anomalous increases observed in ion densities and the sunrise behaviour of conductivity may be an indication of the existence in the mesosphere and stratosphere of such compounds, as they can be photoionized by UV radiation which penetrates to these heights. A detailed study of the ion density variations particularly at low latitudes near sunspot maximum is required to establish the existence of metal hydroxides. The presence of such compounds will have consequences for Sun-weather relationship theories based on electrical conductivity. It has been argued that variations in electrical conductivity affect the electric field between clouds and the Earth and lead to changes in cloud electrification, which

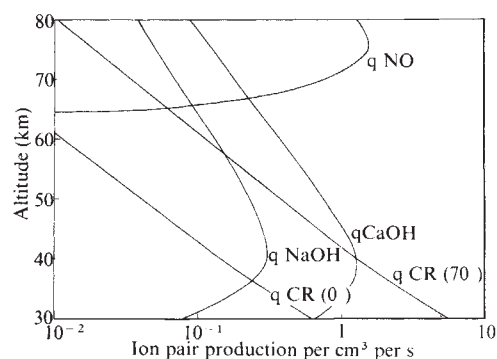


Fig. 2 Ion-pair production rate for the middle atmosphere. Solar zenith angle is 30°. A maximum and minimum cosmic-ray (CR) ionization rate is shown. Top curve is photoionization of NO.

can affect thunderstorm activity and weather¹⁶. If photoionization of meteoric material controls conductivity between 40 and 70 km then solar activity variations of conductivity would be expected. More importantly the ionization rate will be affected not only by solar radiation variability but also by meteor flux.

We thank J. Herman and M. Krauss for useful discussions.

Received 5 March; accepted 29 April 1981.

1. Ferguson, E. C. *Geophys. Res. Lett.* **5**, 1035–1038 (1978).
2. Murad, E., Swider, W. & Benson, S. W. *Nature* **289**, 273–274 (1981).
3. Arijis, E., Nevehans, D. & Ingels, J. *Nature* **288**, 684–686 (1980).
4. Arnold, F. & Henschen, G. *Geophys. Res. Lett.* **8**, 83–86 (1981).
5. Arnold, F., Henschen, G. & Ferguson, E. E. *Planet. Space Sci.* **29**, 185–193 (1981).
6. Hunten, D. M., Turco, R. P. & Toon, O. B. *J. Atmos. Sci.* **37**, 1342–1357 (1980).
7. Rosenstock, H. M., Draxl, K., Steiner, B. W. & Herron, J. T. *J. phys. Chem. Ref. Data* **6**, Suppl. No. 1, (1977).
8. Liu, S. C. & Reid, G. C. *Geophys. Res. Lett.* **6**, 283–286 (1979).
9. Goldberg, R. A. & Aikin, A. C. *Science* **180**, 294–295 (1973).
10. Nicolet, M. & Aikin, A. C. *J. geophys. Res.* **65**, 1469–1485 (1960).
11. Heaps, M. G. *Planet. Space Sci.* **26**, 513–517 (1978).
12. Widdel, H. U., Rose, G. & Borchers, R. *J. geophys. Res.* **81**, 6217–6220 (1976).
13. Smith, D., Adams, N. G. & Church, M. J. *Planet. Space Sci.* **24**, 697–703 (1976).
14. Mitchell, J. D., Sagar, R. S. & Olsen, R. O. *Space Research* Vol. 17 (ed. Rycroft, M. I.) (Pergamon, Oxford, 1976).
15. Croskey, C. L., Hale, L. C. & Leider, S. C. *Space Research* Vol. 17 (ed. Rycroft, M. I.) (Pergamon, Oxford, 1976).
16. Markson, R. & Muir, M. *Science* **208**, 979–990 (1980).

A physicochemical mechanism for the ignition of the Seveso accident

T. G. Theofanous

School of Nuclear Engineering, Purdue University, West Lafayette, Indiana 47907, USA

The so-called Seveso 'explosion' of 10 July 1976 was completely unexpected, and various speculations about its initiation have been entertained, including the possibility of sabotage¹. Seveso has remained the focus of undiminished interest and publicity^{1,2} and the puzzle concerning the occurrence of its accident persists. The explanation proposed here is based on our recent investigations of the relevant chemical kinetics and heat transfer phenomena. This physicochemical mechanism involves the combination of (1) exothermic reactions starting (personal communication from the Givaudan Company) below the previously known 230 °C threshold and (2) a thermal conduction-convection-radiation process responsible for channelling the normally negligible (in global sense) upper reactor wall heat into a small portion (~10%) of the reactor contents, thus yielding the temperature range conducive to these exothermic reactions.

At the time of the accident the Seveso reactor contained 2,030 kg sodium trichlorophenolate, 540 kg sodium chloride, ~1,000 kg ethylene glycol, and 1,600 kg sodium ethylene glycolate, diethylene glycol and polyethylene glycol. Such mixtures are known³ thermally to decompose above 230 °C by virtue of an exothermic chemical reaction yielding gaseous byproducts as well as the highly toxic 2,3,7,8-tetrachlorodibenzodioxin (TCDD). Because such accidents have occurred several times previously, special precautions were taken to prevent the Seveso reactor from reaching this dangerous temperature range. The pressure of the steam utilized for heating (to distill the ethylene glycol) was limited to 12 bar and thus a saturation temperature of 190 °C. Heating was interrupted, and air admitted to re-establish atmospheric pressure, stirring of the reactor contents continued for 15 min and the reactor was left with its liquid contents at the safe temperature level of 158 °C. The accident occurred 7.5 h later.

Early investigations¹ were directed at evaluating the long-term behaviour of reactor-like solutions at various temperatures around the operating level of 158 °C. In particular the potential for slow, slightly exothermic reactions promoted by air (oxidation), or various metal surfaces (catalysis) was thoroughly explored. The results of these studies (indicating stable behaviour) together with recent developments leading to an expectation of local heating (of a small portion of the reactor contents) up to 200–220 °C stimulated interest in a more careful investigation of this temperature range. Differential thermal analysis (DTA) of typical reactor content samples^{4,5} identified two 'slow' exothermic reactions occurring before the decomposition reaction. The first initiates at ~185 °C, peaks at 235 °C and is characterized by an adiabatic temperature rise of 57 °C. The second exotherm initiates at ~255 °C, overlapping with the tail end of the first one, peaks at 265 °C and its adiabatic temperature rise is estimated at 114 °C. The activation energies and the logarithms of the frequency factors are 24,000 K, 29,800 K and 40 and 48 respectively. The decomposition reaction is initiated at ~280–290 °C and exhibits a rapid pressure buildup at ~300 °C. Its adiabatic temperature is estimated at 142 °C.

The adiabatic induction times are 2.1 h and 30 min for the first (~200 °C) and second (~240 °C) exotherm respectively. It is easy to see how the decomposition reaction may be reached quite rapidly (compared with the Seveso time scale) in an adiabatic fluid volume that has been heated to ~200 °C. However, the actual thermal process driving the ignition is approximated well by the air-liquid interface held at a constant temperature in the 200–220 °C range. In such conditions the temperature gradients may be steep, the heat losses may be substantial and the thermal ignition theory⁶ would predict times for self-accelerating conditions substantially longer than the adiabatic induction times (roughly 10 times longer). However, in relation to this theory the actual situation is complicated by (1) three consecutive, and partly overlapping reactions (the first two, for which reactant depletion cannot be neglected) serve for warmup; and (2) uncommon and complication-yielding boundary conditions during the warmup period (that is, assuming the surface temperature to a constant value during this period accentuates heat losses and produces overestimation of ignition times). We must regard, therefore, the thermal ignition theory results only as qualitative estimates indicating that: (1) heatup of the liquid surface to ~200–220 °C represents a credible initiation of the decomposition-yielding chain of reactions within ~7 h; and (2) heat losses and reactant depletion need be taken into account for a more precise evaluation of the warmup

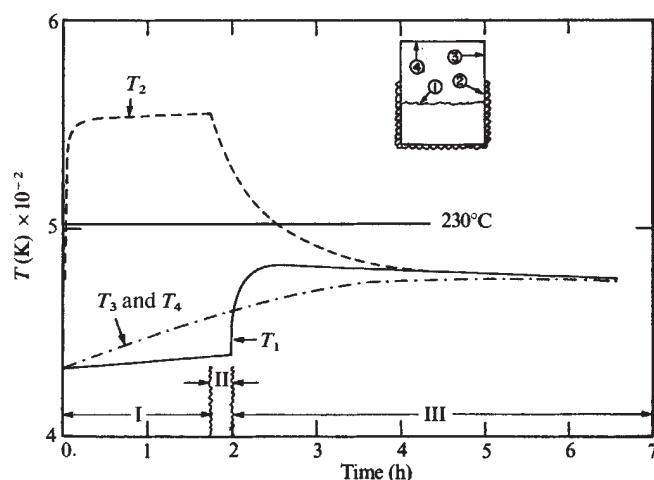


Fig. 1 Calculated thermal transient of the Seveso reactor walls and reaction mixture free surface. Inset is schematic of the Seveso reactor. Surface 2 corresponds to the unsubmerged heated surface (backed by steam coils). Emissivities of 1 and 0.5 were used for the liquid and stainless wall respectively.

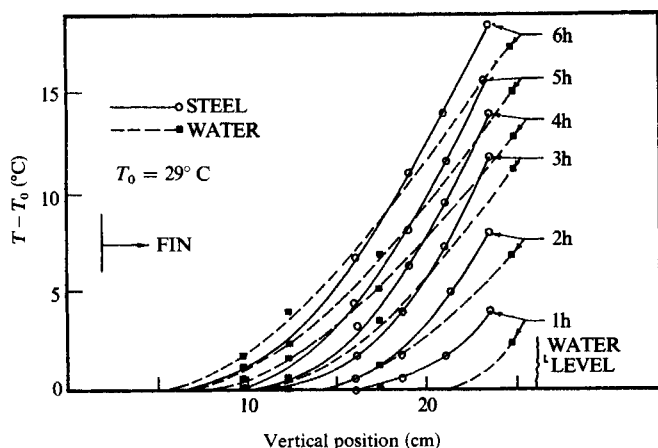


Fig. 2 Experimental demonstration of the stratification phenomena in the submerged fin case with a driving temperature gradient of $\sim 10^\circ\text{C cm}^{-1}$. Lateral temperature gradients in the liquid were negligible.

period. The thermal-physical processes driving the initial stages of the ignition are now considered.

The geometry of the reactor, the heating jacket and the liquid level are indicated in Fig. 1 (inset). Distillation of ethylene glycol was carried out under vacuum (10 torr) by means of 12 bar superheated ($\sim 300^\circ\text{C}$) steam. At the time that heating was interrupted the distillation had proceeded for $1\frac{1}{2}$ h (period I). During this period surface 2 (I_2) heats quickly to a temperature near that of the steam. The liquid remains at a uniform temperature of $\sim 160^\circ\text{C}$ and surfaces 3 and 4 (I_3 and I_4) experience slow heating due to thermal radiation from surface 2. During the subsequent 15 min (period II) with continuing agitation the liquid remains at the uniform temperature of $\sim 160^\circ\text{C}$, the submerged reactor wall cools also to this same temperature, surfaces 3 and 4 continue to heat up slowly while surface 2 rapidly cools by radiation to all three other surfaces. Clearly in the neighbourhood of surface junctions steep temperature gradients and additional cooling will exist; however, a reasonable estimate of temperature transients may be obtained neglecting temporarily such gradients. Finally, when stirring is also interrupted (period III) the liquid surface will quickly heat up to an 'equilibrium' temperature as surface 2 continues to cool and surfaces 3 and 4 continue to heat till all four surfaces come to equilibrium. A straightforward lumped parameter model was devised⁷ to describe this sequence of events quantitatively. The four unsteady-state energy balances include heat losses to environment, steam heating of surface 2 and a diffuse-grey formulation for radiative exchange among all four surfaces. During period III the liquid surface temperature is evaluated by assuming a conduction controlled the thermal boundary layer. Note that during distillation the vapour pressure is too low to produce significant inter-surface shielding. Similarly, during periods II and III the air presence in the gas space precludes such interference. The system was integrated numerically and the results of a representative calculation are shown in Fig. 1. Note that at the inception of period III only a small portion, if any (as actual cessation of all liquid motions, as assumed in the analysis, will occur in actuality well into period III), of the wall is around the threshold of 230°C , while the liquid surface temperature remains well below this limit. In fact, the results are not very sensitive to parameter variations (such as the wall emissivity) always yielding surface liquid temperatures in the $200\text{--}220^\circ\text{C}$ range.

Figure 1 shows that the low liquid thermal diffusivity yields an isolation effect, whereby the liquid surface heats to nearly an equilibrium value consistent with transmitting a very low-energy flux. On the other hand, radiative exchange within surface 2 will tend to sharpen the temperature gradient near its boundary with surface 1 (I_1). The stainless wall has a thermal diffusivity two orders of magnitude higher than that of the liquid. One would

expect significant conduction along the wall, creation of horizontal density gradients in the liquid and establishment of convective currents⁸. There is no direct information published about this complex heat transfer process but large-scale circulation currents are expected. These currents could destroy the thermal stratification picture developed above yielding even larger margins of maximum liquid temperatures from the 230°C limit. We will argue instead that this convection-conduction phenomenon acts in concert with the radiation process to produce an extremely localized hot stratum at the liquid surface.

As a simple first-order analysis we may view the vessel wall as a submerged fin. If T_w is the temperature of the fin base (at the liquid surface) and k_w the wall thermal conductivity, a quasi-static analysis indicates that the energy dissipated, by natural convection in a fluid of temperature T_∞ , is given by $q = (Nu_B k_b / k_w)^{1/2} k_w (T_w - T_\infty) / B$. The Nusselt number is based on the fluid thermal conductivity k_b and fin thickness B . For the present situation it is easy to see that the square-root term is of the order of one. The fin-internal temperature gradient is, therefore, confined within a region of order B from the fin base. The same conclusion is also obtained for the case of a stratified liquid layer, that is with an imposed vertical temperature gradient in the bulk. An experimental confirmation of this trend was also obtained. A portion of metal (carbon steel) slab was immersed into a liquid (water) filled container. A temperature gradient in the well insulated non-submerged portion was established by steadily heating at one end and conducting into the liquid at the other. Determination of the resultant temperature fields was made by thermocouple measurements at a large number of points. One set of typical results is presented in Fig. 2. In particular note the coincidence of the fin and surrounding medium thermal boundary layers. Evaluation of growth histories of these thermal boundary layers indicates a process controlled by the liquid thermal diffusivity, that is $\delta \sim (\alpha_1 t)^{1/2}$.

The combined effects of radiation and fin processes were also studied in an experiment more closely simulating the Seveso reactor using a cylindrical stainless-steel vessel, 76.2 cm in diameter and 122 cm tall. The wall thickness was selected at 0.277 cm such that the liquid surface area to wall cross-section area will approximate the Seveso reactor value of 60. This vessel was uniformly wrapped with cable heater and insulated to a degree approximating that of the Seveso reactor. The vessel was uniformly heated, while empty, to $\sim 80^\circ\text{C}$ above room temperature (see Fig. 1 at beginning of period III). Room-temperature ethylene glycol was quietly admitted through a line

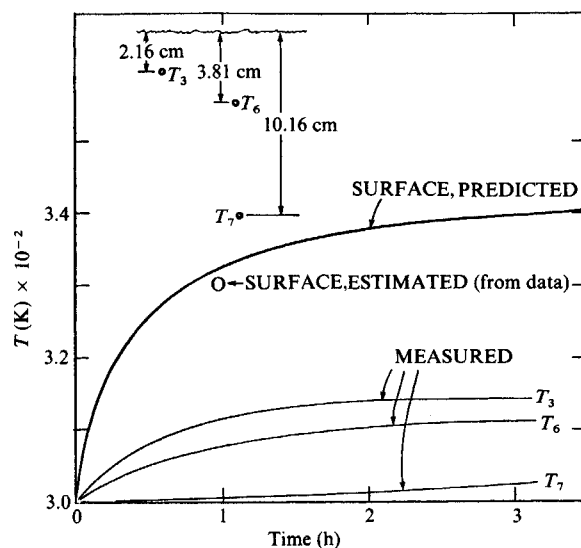


Fig. 3 Measured temperature transients near the liquid surface as indicated. Experimental and analytical simulation of the thermal transient in the Seveso reactor. The prediction is based on the lumped parameter radiation/diffusion model (emissivity $\epsilon = 0.25$).

at the vessel base to a height of ~ 30 cm. The system was left at rest and temperature transients at the vessel walls and liquid contents were recorded with thermocouples. Due to space limitations only the essence of the results as summarized in Fig. 3 may be discussed here. The experimental data clearly show the downward development of the thermal boundary layer and they are in quantitative agreement with the $(\alpha_1 t)^{1/2}$ dependence. The spatial dependence of these profiles also agrees with that predicted for transient conduction into a semi-infinite domain with a constant temperature of the surface. Hence, this surface temperature may be estimated and was found to be in good agreement with our theoretical predictions as shown in Fig. 3. A rapid liquid surface temperature increase of 30–40 °C similar both in trend as well as in magnitude to that calculated for the Seveso reactor (Fig. 1) is thus deduced demonstrating the nature of the stratification phenomena and the simulation capability of our experimental setup.

I thank Dr M. Saito for experimental assistance

Received 9 March; accepted 14 April 1981.

- Hay, A. *Disasters* (to be published).
- Hay, A. *Nature* **262**, 636 (1976); **267**, 384 (1977); **273**, 583 (1978).
- Milnes, M. H. *Nature* **232**, 395 (1971).
- Salomon, Ch. M. *Bericht über den Störfall in Seveso, Sonderkolloquium am 4* (Dechema-Haus, Frankfurt am Main, 1980).
- Salomon, Ch. M. *Meet. Industrial Health and Safety Group of Soc. Chem. Ind. London* (1981).
- Merzalanov, A. G. & Arerson, A. E. *Combust. Flame* **16**, 89 (1971).
- Theofanous, T. G. *Meet. Industrial Health and Safety Group of Soc. Chem. Ind. London* (1981).
- Theofanous, T. G. *Bericht über den Störfall in Seveso, Sonderkolloquium am 4* (Dechema-Haus, Frankfurt am Main, 1980).

Proterozoic magnetic overprinting of Archaean rocks in the Canadian Superior Province

Larry D. Schutts & David J. Dunlop

Department of Geology and Geophysics Division, Department of Physics, University of Toronto, Toronto, Canada M5S 1A7

We have measured the various components of natural remanent magnetization (NRM) exhibited by mafic extrusive and intrusive rocks from the Abitibi Subprovince of the Canadian Precambrian Shield. In addition to providing what seems to be a primary Archaean pole for the region, the results testify to three episodes of magnetic 'overprinting' at approximately 2,600, 2,100 and 1,800 Myr, respectively. The two earlier episodes coincide with the intrusion of Matachewan and Abitibi diabase dyke swarms, here and elsewhere in the Superior Province. Magnetic overprinting, in these cases, extends far beyond areas of surficial dyke exposure. The third episode postdates magmatic activity in the Abitibi region, but coincides in age with the 1,900–1,700 Myr Hudsonian orogeny in the Churchill and Bear Provinces and the Penokean orogeny in the Southern Province. Although magnetic overprints presumed to be of this age were known previously from Abitibi dykes¹ and Nipissing sills^{2–4}, our data are the first evidence of a 1,800 Myr magnetic overprinting on a regional scale in the Superior Province.

Palaeomagnetism is a powerful method of detecting and dating regional geological events in metamorphic terrains, particularly in Precambrian regions where complex structures, deep erosion, and various generations of burial and intrusion often confuse relations. Each geological event can result in thermal and/or chemical resetting of part of the pre-existing NRM. In favourable circumstances, careful multicomponent palaeomagnetic analysis can isolate vectorially superimposed NRM components corresponding to primary magnetization and one or more generations of NRM overprinting^{5–8}. The geologi-

cal events responsible are dated by comparing the palaeomagnetic pole corresponding to each NRM component with an established apparent polar wander path (APWP) for the continent or continental block.

The Abitibi greenstone belt of the Superior Province is particularly suited to a regional multicomponent palaeomagnetic study. The late Archaean and early Proterozoic units have been thoroughly mapped and the Laurentian APWP for this time interval is reasonably well established (Tracks 4, 5 and 6 of refs 9, 10). The Abitibi belt is characterized by mafic to felsic volcanic piles with coeval intrusions, volcanogenic sediments and intruding granitic batholiths^{11,12}. In the Kirkland Lake, Ontario area, two such Archaean piles, plus the younger portion of a third pile, are preserved in a large east-plunging synclinorium¹³. The gabbroic intrusive bodies occur as both conformable sills¹⁴ and as stocks, which seem to have been feeders for some of the flows¹⁵. Granitic bodies, apparently representing a later Archaean event, intrude the mafics on the north, west and south as batholiths and within the synclinorium as occasional stocks and plugs. Magmatic activity continued with intrusion of the north-south trending Matachewan diabase dyke swarm and essentially concluded with intrusion of the north-east trending Abitibi diabase dykes.

Archaean activity in the area was evidently episodic in occurrence. The youngest volcanic pile, which represents more than 70% of the 50,000 m total pile thickness¹³, is bracketed in age by $2,710 \pm 2$ Myr and $2,703 \pm 2$ Myr U–Pb zircon dates¹⁶. These figures compare with $2,725 \pm 2$ Myr and $2,703 \pm 3$ Myr zircon ages bracketing the correlative sequence near Timmins, some 120 km to the west^{17,18}. Granite ages are less narrowly defined. The oldest dates, determined by both K/Ar and Rb/Sr methods, fall within the 2,500–2,600 Myr range for the Round Lake Batholith^{19,20}. The granites are cut by Matachewan dykes which, with a $\lambda(^{87}\text{Rb})$ of $1.42 \times 10^{-11} \text{ yr}^{-1}$ (ref. 21), yield Rb/Sr isochron data implying an age of $2,633 \pm 93$ Myr (ref. 22). The relative age relation, however, is clear. The Matachewan dykes are seen to intrude the granite in many locations and absolute ages for both units probably fall within the 2,600–2,700 Myr span. The time of Abitibi dyke intrusion is more precisely defined and implies a discrete early Proterozoic event at 2,150 Myr (refs 22, 23).

For our study, 157 block samples were taken over an areal extent of 5,200 km² in the Kirkland Lake Matheson area. The suite of samples includes basalts and gabbros from the youngest synclinorium sequence, as well as diabases from 15 Matachewan and three Abitibi dykes. Several 2.5-cm cylindrical specimens were cut from each sample and subjected to detailed stepwise thermal (to 700 °C) and alternating field (to 100 mT or 1,000 Oe) demagnetization, using standard commercial demagnetizers. NRM intensities and directions were measured on a computerized commercial spinner magnetometer, which provided a minimum detectable signal of 10^{-4} A m^{-1} ($10^{-7} \text{ e.m.u. cm}^{-3}$) and on-line statistical analysis of results. (A more detailed analysis of the alternating field and thermal demagnetization results, and the results of experiments in progress designed to characterize more fully the magnetic remanence carriers, will be published elsewhere²⁴.)

Except for a few basalt and gabbro samples, in which an intermediate (~ 350 °C) Curie temperature carrier of magnetization was indicated, all samples exhibited maximum thermal demagnetization blocking temperatures near 580 °C, the Curie temperature of pure (titanium-free) magnetite. Basalt, gabbro and Abitibi dyke blocking temperatures tended to be distributed from 100 or 200 °C to 580 °C, while Matachewan dyke blocking temperatures were generally concentrated just below 580 °C. This evidence of relatively pristine magnetite as the principal NRM carrier is in accord with petrographical observations from this and other studies^{15,25} implying only very low-grade regional metamorphism in the area.

Different methods were used to analyse directional changes during demagnetization. Stable endpoints, Briden Index criteria²⁶, Zijdeveld-type plots²⁷, vector subtraction diagrams²⁸

Table 1 Summary of site mean palaeomagnetic results

	<i>D</i>	<i>I</i>	α_{95}	<i>k</i>	<i>R</i>	<i>N</i>	Palaeopole	
Characteristic directions								
Archaean basalt-gabbro	114°	-32°	11°	15	12.2	13	29° S	4° E
Matachewan dykes and baked contacts	21°	+13°	6°	44	13.7	14	45° N	70° E
Abitibi dykes and baked contacts	277°	+70°	8°	52	7.9	8	41° N	230° E
Secondary directions								
Archaean basalt-gabbro Matachewan-like	28°	+25°	11°	19	9.5	10	48° N	57° E
Archaean basalt-gabbro Abitibi-like	298°	+77°	9°	30	10.7	11	54° N	241° E
All units (sample averaged) post-Abitibi	166°	+56°	8°	14	26.1	28	4° S	291° E

D, *I* are declination and inclination of magnetization; α_{95} , *k* and *R* are Fisher statistical parameters; *N* is the number of sites averaged.

and converging remagnetization circles²⁹ all provided data relevant to the segregation of magnetic components. The isolated components were classified as either characteristic or secondary remanences. We define the 'characteristic NRM' in a multicomponent sum to be that component possessing the higher coercivities and blocking temperatures. A 'secondary NRM' is that component of lower coercivity and blocking temperature which overprints a characteristic NRM. Because our results imply that the reset magnetizations are of thermal as opposed to chemical origin, this scheme seems justified.

A standard three-tier (specimen/sample/site) analysis of characteristic and secondary NRM directions was carried out, a site being defined as a single stock, dyke, sill or flow and its baked contacts. Site-mean average magnetization directions are listed in Table 1 and plotted in Fig. 1. All NRM directions listed in Table 1 exhibit both normal and reverse polarities. Palaeopoles corresponding to each magnetization direction are shown in Fig. 2 and compared with the Tracks 5 and 6 of the North American APWP from refs 9, 10. Archaean ages have been updated to agree with the recent radiometric dates for the volcanics.

Abitibi and Matachewan dyke characteristic directions are well-defined and closely agree with the results of previous studies^{1,10,30,31}. Evidence that these directions record the geomagnetic field direction at the time of dyke intrusion is provided by detailed Abitibi-Matachewan and Matachewan-gabbro contact tests in Munro Township. The pattern is the same in either case. Near the contact, NRM directions in the older unit have been reset parallel to those of the intrusive body. Well away from the contact, NRM directions in the older unit approach the characteristic direction. In an intermediate 'hybrid' zone, both NRMs are superimposed, the resultant direction swinging towards the older characteristic during

magnetic cleaning. The maximum blocking temperature range of the overprinted NRM decreases in a regular fashion with distance from the contact. These data are strong evidence for partial to complete thermal resetting of primary NRM in the country rock at the time of dyke intrusion.

Basalt and gabbro samples were taken from the upper volcanic-plutonic pile of the synclinorium. The suite contains blocks from the Kinojevis tholeiite Group and the underlying Stoughton-Roquemaure and Larder Lake komatiite Groups. Three separate magnetic components, common to rocks from all three units, were isolated. Two of these components have the Abitibi and Matachewan dyke characteristic directions (Fig. 1*b*) and we believe them to be regional thermal overprints. The third component is similar to one found in other Superior Province Archaean units^{10,32} and possibly records the geomagnetic field at the time these units were emplaced.

Several lines of evidence imply that the Abitibi-like and Matachewan-like components are secondary thermal overprints. They primarily affect the lower coercivity and blocking temperature spectra of the samples exhibiting them. They do persist throughout these spectra in some instances, but samples displaying this behaviour generally have coercivities and blocking temperatures centred about low absolute values. The directional changes with increasing demagnetization level are invariably in the sense suggested by the relative age data. Matachewan-like directions swing towards the older Archaean characteristic direction. Abitibi-like directions move towards the same Archaean direction or towards a Matachewan direction. These swings occur with both alternating field and thermal demagnetization. As both secondary and characteristic remanent magnetizations are lost by ~580 °C, the possibility of chemical overprinting seems remote. Partial thermal resetting of NRM as a result of, first, Matachewan and then Abitibi intrusion

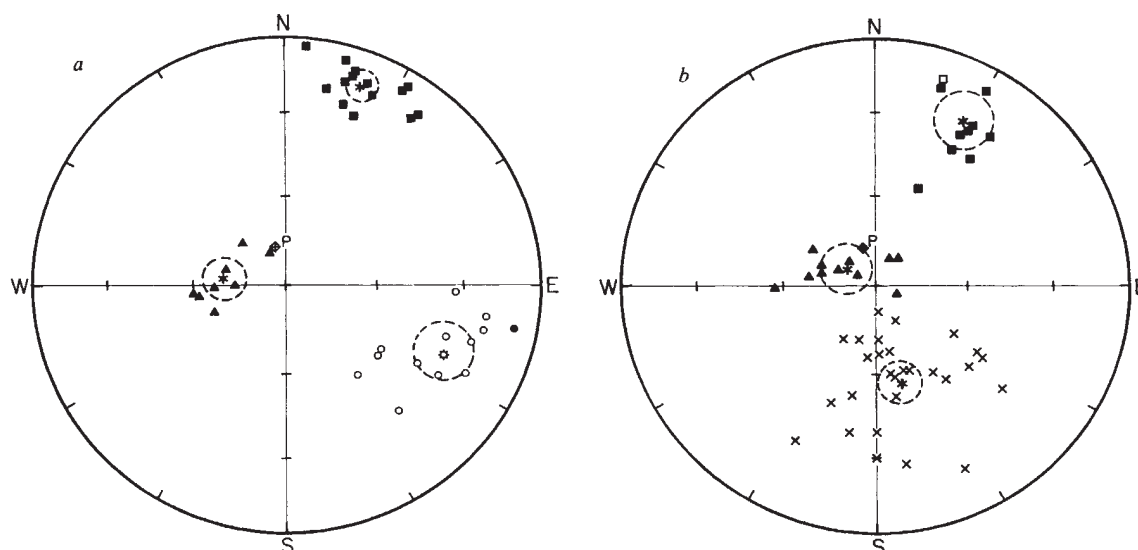


Fig. 1 Abitibi Subprovince site mean magnetization directions after optimum AF and/or thermal cleaning. Solid (open) symbols denote positive (negative) inclination. P (⊕) is the present Earth's field direction. ×, Average of site mean directions. *a*, Characteristic directions: ○, basalt-gabbro; ■, Matachewan diabase; ▲, Abitibi diabase. *b*, Secondary directions: ■, Matachewan-like; ▲, Abitibi-like; ×, Post-Abitibi (sample directions).

seems to be the most likely explanation for the occurrence of both secondary components. As the overprints are generally found in areas containing little or no surficial expression of such dyke intrusion, the entire region was probably reheated, although the existence of dykes which do not break the present surface is also possible.

Averages of the basalt and the gabbro site-mean characteristic directions are nearly identical. Hence, results for both rock types are combined to give a 'basalt-gabbro' characteristic direction. Throughout the area sampled, the gabbros occur primarily as bodies intruding the highly folded basalts. The basalt directions, corrected for structure, become essentially random and a case might be made for regional resetting at the time of gabbro intrusion. The common characteristic direction, however, is also found in basalts far removed from any obvious intrusion. It is also found in a conformable gabbroic sill within the volcanics, which yields a negative fold test (McCool Township). These observations, combined with geological suggestions that the volcanics and plutonics are consanguineous¹⁵ and that the structure in the volcanics developed during deposition³³, provide strong evidence for the primary nature of the 'basalt-gabbro' characteristic.

This scheme does not take into account the ubiquitous granite intrusions. They apparently represent an Archaean event which postdates synclinorium mafic activity, but predates Matachewan dyke intrusion. Current palaeomagnetic evidence does not rule out the basalt-gabbro characteristic direction being the result of a regional granitic intrusion heating. Low granite intrusion temperatures and gravity data suggesting that the larger bodies are limited to the synclinorium periphery^{11,34} oppose this possibility. Ultimate resolution of the question, however, awaits detailed palaeomagnetic study of the granitic units.

Although plutonic and volcanic activity in the area essentially ceased with intrusion of the Abitibi dykes, a post-Abitibi, thermally-reset NRM component has been isolated in the lower coercivity-blocking temperature ranges of 28 of the samples studied. These components partially overprint each of the other characteristic and secondary directions already discussed and show no preference for rock type or geographic locality. They are, however, encountered only sporadically throughout the various sites. As a result, sample means have been used to calculate the average direction and palaeopole.

Poles corresponding to these sample-mean directions are grouped around the 1,700–1,900 Myr segment of Tracks 4 and 5 (Fig. 2). Such components have previously been found in both Abitibi dyke¹ and Nipissing sill^{2,3} samples from the region and have been assigned secondary magnetization status and late Aphebian age on the basis of palaeomagnetic work in the Cobalt area⁴. By using the ⁴⁰Ar/³⁹Ar step-heating variant of K/Ar dating, Hanes and York have uncovered evidence for a late Aphebian heating event in Munro Township²³. Samples taken from an Abitibi dyke yielded a 2,150 Myr age for the dyke interior and 1,700–1,800 Myr ages for the edge zone contact with a Matachewan dyke. Several samples from this Matachewan dyke and one from the country rock gabbro also gave 1,700–1,800 Myr plateaus. The detailed Abitibi-Matachewan contact test made for this study was conducted where the same Abitibi dyke investigated by Hanes and York cut a Matachewan dyke 150 m east of the one they sampled. The post-Abitibi secondary direction was found in both units, most frequently near the contact. A late Aphebian age for this post-Abitibi NRM component is strongly implied.

More than half of the post-Abitibi directions discovered occurred in samples taken near intrusive contacts or zones of some faulting. A high heat flow in the region 1,700–1,900 Myr ago, facilitated by hot fluid transport through faults and intrusive contact zones of weakness, could explain both magnetic and argon resetting. A re-echoing of the Southern Province Penokean orogeny, for example, might be postulated as a source for the high heat flow. Although late Aphebian chemical resetting of remanence in some samples cannot be ruled out, Curie temperatures and preliminary microscopic examinations in

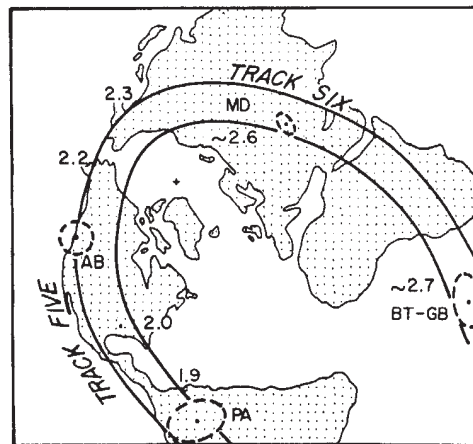


Fig. 2 Palaeomagnetic poles corresponding to Fig. 1 NRM directions as compared with the Laurentian APWP for the interval 1,900–2,700 Myr. BT-GB, basalt-gabbro characteristic. MD, Matachewan diabase characteristic—secondary. AB, Abitibi diabase characteristic—secondary. PA, post-Abitibi secondary.

reflected light suggest that thermal resetting should be the favoured hypothesis. Single Curie temperatures in the range 570–580 °C and intermediate deuteric alteration states with little sign of maghaemization or haematization of the titanomagnetite phase imply no significant metasomatic alteration of the remanence carrier since formation.

Alternatively, the resetting may not be related to faulting and intrusive contacts on a regional scale and a late Aphebian unroofing event might be invoked to explain the younger ages and directions. The region was covered by Huronian sediments at ~2,300 Myr (ref. 35). A Schwarz-type^{36,37} depth of burial calculation, using minimum blocking temperatures of the pre-overprint component, has been made for samples from the Munro Township Abitibi-Matachewan contact. The results imply an ambient temperature for the present erosional surface of 190 °C at the time of Abitibi dyke intrusion. Assuming a 25 °C km⁻¹ geothermal gradient, this temperature implies a 7–8 km depth of burial at 2,150 Myr. The area was still covered, then, during the middle Aphebian, but a late Aphebian unroofing may certainly be postulated.

The close agreement of our independently dated post-Abitibi palaeopole for the Superior Province with palaeopoles ~1,800 Myr in age from the Churchill and Bear Provinces has broad implications for the tectonic history of Laurentia. This evidence implies that the Churchill orogen and the adjacent, relatively stable Superior craton occupied their present relative positions during the 1,900–1,700 Myr Hudsonian orogeny. The orogeny, therefore, may not have been the result of large-scale relative motions between these two blocks, such as are involved in modern day plate collision and subduction. The Hudsonian orogeny may represent an essentially ensialic reactivation of the Archaean Laurentian craton that incorporated both the Superior and Churchill Provinces.

Our preliminary Abitibi Subprovince results demonstrate that the geological history of complex Archaean terrains can be clarified with the help of palaeomagnetism. Resolved multi-component magnetizations can provide, in effect, a capsule history of thermal events throughout an entire region.

This research has been supported by Natural Sciences and Engineering Research Council Canada and Energy Mines and Resources Canada Research Grants to D. J. D. and by University of Toronto and Ontario Graduate Fellowships to L. D. S.

Received 22 December 1980; accepted 8 April 1981.

1. Laroche, A. *Can. J. Earth Sci.* **3**, 671 (1966).
2. Patel, J. P. & Palmer, H. C. *Can. J. Earth Sci.* **11**, 353 (1974).
3. Symons, D. T. A. & Londry, J. W. *Can. J. Earth Sci.* **12**, 940 (1975).
4. Roy, J. L. Lapointe, P. L. *Can. J. Earth Sci.* **13**, 749 (1976).
5. Buchan, K. L., Berger, G. W., McWilliams, M. O., York, D. & Dunlop, D. J. *J. Geomagn. Geoelectr.* **29**, 401 (1977).
6. McElhinny, M. W. & Senanayake, W. E. *J. geophys. Res.* **85**, 3523 (1980).
7. Beckmann, G. E. J. & Mitchell, J. G. *Earth planet. Sci. Lett.* **30**, 269 (1976).

8. Morgan, G. E. *Nature* **259**, 382 (1976).
9. Irving, E. & McGlynn, J. C. *Trans. R. Soc. A* **280**, 433 (1976).
10. Irving, E. & Nalder, A. J. *J. Geol.* **85**, 157 (1977).
11. Goodwin, A. M. & Ridler, R. H. *Geol. Surv. Canad. Pap.* 70-40, 1 (1970).
12. Goodwin, A. M. *Geol. Ass. Can. Spec. Pap.* **16**, 205 (1977).
13. Jensen, L. S. *Ont. Geol. Surv. Misc. Pap.* No. 90, 64 (1979).
14. Pyke, D. R., Nalder, A. J. & Eckstrand, O. R. *Bull. geol. Soc. Am.* **84**, 955 (1973).
15. Jensen, L. S. *Ont. Geol. Surv. Rep.* 165 (1978).
16. Nunes, P. D. & Jensen, L. S. *Ont. Geol. Surv. Misc. Pap.* 92 (1980).
17. Pyke, D. R. *Ont. Geol. Surv. Misc. Pap.* No. 82, 73 (1978).
18. Nunes, P. D. & Pyke, D. R. *Ont. Geol. Surv. Misc. Pap.* 92 (1980).
19. Aldrich, L. T. & Wetherill, G. W. *J. geophys. Res.* **65**, 337 (1960).
20. Purdy, J. W. & York, D. C. *Can. J. Earth Sci.* **5**, 699 (1968).
21. Steiger, R. H. & Jäger, E. *Earth planet. Sci. Lett.* **36**, 359 (1977).
22. Gates, T. M. & Hurley, P. M. *Can. J. Earth Sci.* **10**, 900 (1973).
23. Hanes, J. A. & York, D. C. *Can. J. Earth Sci.* **16**, 1060 (1979).
24. Schutts, L. D. (in preparation).
25. Jolly, W. T. *Can. Miner.* **12**, 499 (1974).
26. Briden, J. C. *J. geophys. Res.* **77**, 1401 (1972).
27. Zijdeveld, J. D. A. in *Methods of Paleomagnetism* (eds Collinson, D. S., Creer, K. M. & Runcorn, S. K.) 172 (Elsevier, Amsterdam, 1967).
28. Halls, H. C. *Earth planet. Sci. Lett.* **43**, 303 (1979).
29. Halls, H. C. *Phys. Earth planet. Inter.* **16**, 1 (1978).
30. Strangway, D. W. *J. Geol.* **72**, 648 (1964).
31. Fahrig, W. F., Gaucher, E. H. & Larochelle, A. *Can. J. Earth Sci.* **2**, 278 (1965).
32. Dunlop, D. J. *Phil. Trans. R. Soc. (in the press)*.
33. Jensen, L. S. *Toronto '78 Field Trips Guidebook*, 237 (Geological Association of Canada, 1978).
34. Middleton, R. S. *Ont. Div. Mines, Geosci. Rep.* 135 (1976).
35. Fairbairn, H. W., Hurley, P. M., Card, K. D. & Knight, C. J. *Can. J. Earth Sci.* **6**, 489 (1969).
36. Schwarz, E. J. *Can. J. Earth Sci.* **14**, 82 (1977).
37. Buchan, K. L. & Schwarz, E. J. *Can. J. Earth Sci.* (submitted).

Low-angle normal faults in the Basin and Range Province: nappe tectonics in an extending orogen

Brian Wernicke

Department of Earth and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Cenozoic normal fault mosaics bounded beneath by a basal fault in the Basin and Range Province (BRP) have traditionally been described either in terms of large-scale surficial gravity sliding or by some form of *in situ* lower plate accommodation. I suggest here that these areas may be an extensional analogue to thin-skinned compressional orogens, a process which may even dominate active BRP tectonics.

These terranes consist of a mosaic of rock sheets bounded by high- and low-angle normal faults^{1,2} which resemble (and have been misinterpreted as) compressional nappe piles. In many areas, the normal faults 'bottom out' into a single, subhorizontal fault zone below which either minor or no tectonism has occurred³⁻⁶, or a roughly synchronous, more ductile style of deformation is present^{7,8}. The minimum areal extent of the basal faults is typically measured in thousands of square kilometres.

For example, in east-central Nevada, normal fault systems developed in Palaeozoic miogeoclinal strata, Mesozoic and Tertiary intrusives, and Tertiary volcanics end abruptly at a basal fault (Snake Range decollement³, Fig. 1) accompanied by a thin (0–100 m) zone of tectonite marble. Offsets on the basal and higher faults are difficult to constrain, but Nelson reports a 20-km minimum horizontal offset for one fault alone in the Deep Creek Range⁹. My palinspastic reconstruction of the Northern Egan and Cherry Creek ranges (Fig. 2) indicates a 16.5-km offset for one sheet in the mosaic, and the per cent increase in original width of the area is ~300%. As the basal and higher faults in Fig. 1 are commonly subparallel to bedding and often cut out thousands of metres of section, total offsets are probably at least in the range of several tens of kilometres. Most authors except one¹⁰ who have mapped in the area outlined in Fig. 1 agree that the transport direction of the allochthonous sheets was eastwards. The faults also tend to cut consistently downsection to the east^{11,12}, (Fig. 2). This terrane apparently has every attribute of a thin-skinned compressional orogen (basal decollement, far-travelled allochthons, consistent direction of section transgression and transport) except (1) the low-angle faults habitually omit rather than repeat section, (2) associated high-angle faults are predominantly normal instead of reverse, and (3) folding is of subdued importance¹³. Though early workers attempted to correlate these faults with Mesozoic thrusting in the Sevier belt to the east³, there is considerable evidence to support a Tertiary, extension-related origin for most of the deformation¹³⁻¹⁵.

Models for these areas must account for three observations: (1) the basal and higher faults predominantly represent extension; (2) the rocks below the basal faults are tectonically inert during transport of the allochthonous sheets and high-angle fault blocks; and (3) the areal extent and transport distances of the allochthons are large. Current models fall into two categories. The first is a localized gravity-slide model in which domal upwarps shed their cover in a megalandslide fashion¹⁵⁻¹⁸ (Fig. 3a). The second invokes penetrative ductile stretching and/or igneous intrusion to drive extension in the immediately overlying brittle fault mosaic^{2,7,8,19,20} (Fig. 3b). Landslide models are weakened by the scarcity of geometrically required areas of shortening correlative in age, size and direction of transport with the extensional terranes. Thus, the model agrees with observations (2) and (3), but not with observation (1). *In situ* ductile stretching or intrusion models lose appeal when it is noted that most metamorphic fabrics and igneous bodies (if any) beneath the basal faults are of inappropriate age or geometry to be rigorously linked to normal faulting^{5,6,21,22}. Thus, the model is in harmony with observations (1) and (3), but not with observation (2).

The key to the problem may lie in our understanding of compressional orogens, in which the horizontal movement of thin sheets of rock above an underthrusting slab may occur many

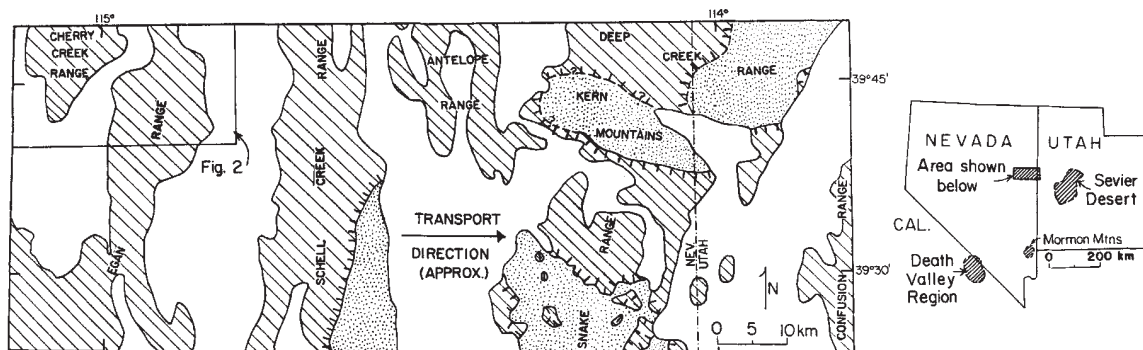


Fig. 1 Location map and tectonic interpretation of east-central Nevada, showing autochthon (stippled), basal fault zone (line with tick marks, queried where uncertain) and normal fault mosaic (hatched). Basal fault in the Schell Creek Range is locally tilted steeply westwards, implying continued deformation after it ceased moving, possibly on a lower basal fault. Tilting and faulting in the Northern Egan Range may be related to this second event rather than to movement on the basal fault shown here.

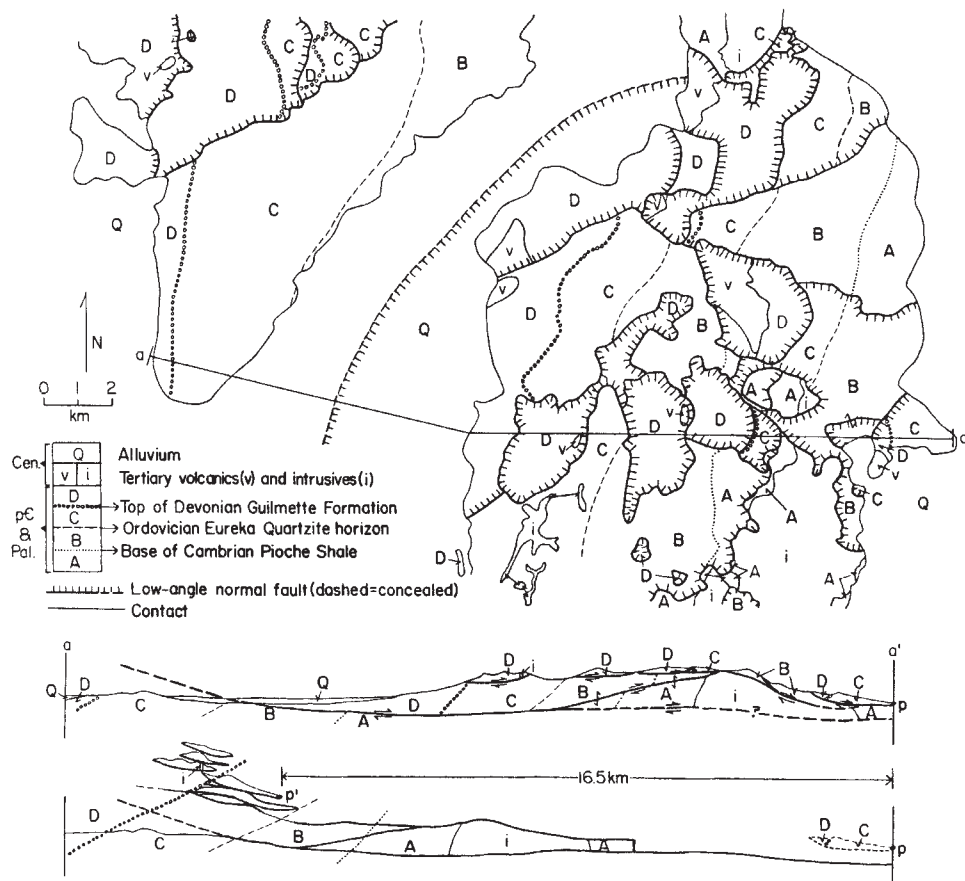


Fig. 2 Highly simplified map, cross-section and palinspastic reconstruction of the area indexed in Fig. 1. Reconstruction does not remove tilting, which was probably synchronous with faulting. Armstrong¹⁵ interpreted the sheets as having formed exclusively by high-angle normal faulting, then subsequently rotated to the horizontal. Geology by Fritz¹¹.

tens (even hundreds) of kilometres from the nearest possible site of coeval lower plate shortening. Applying this concept to extensional tectonics to explain relationships observed in the BRP, I suggest that the basal faults 'root' into the crust (and perhaps the mantle) at a low angle, thereby serving as a narrow zone of decoupling between two thin sheets of rock (Fig. 3c). Towards the thick end of the upper plate, extensional faulting may be negligible or absent, but towards the thin end it loses its ability to remain coherent and becomes a thin-skinned extensional fault terrane. Lower plate extension may therefore be far removed horizontally (perhaps a distance many times the areal width of upper plate normal faulting) and at much greater depth

than the thin-skinned extensional belt. This hypothesis allows for a rigid autochthon and eliminates the need for immense terranes of coeval shortening, thus satisfying all three observations.

As shown in Fig. 3c, rock sheets (nappes²³) may be carried along the basal fault and accreted to the autochthon, kinematically similar to nappe emplacement along the soles of overthrust masses in compressional belts. Note that the nappes in Fig. 3c have converged on one another, a relationship that superficially resembles shortening. However, the amount of movement between them is actually a measure of crustal divergence. If the faults had developed in a pristine sedimentary

Fig. 3 First-order kinematic schemes to generate low-angle normal fault mosaics above abasal fault. *a*, Megalandslide model; *b*, *in situ* ductile stretching or intrusion model; *c*, rooted, low-angle normal fault model.

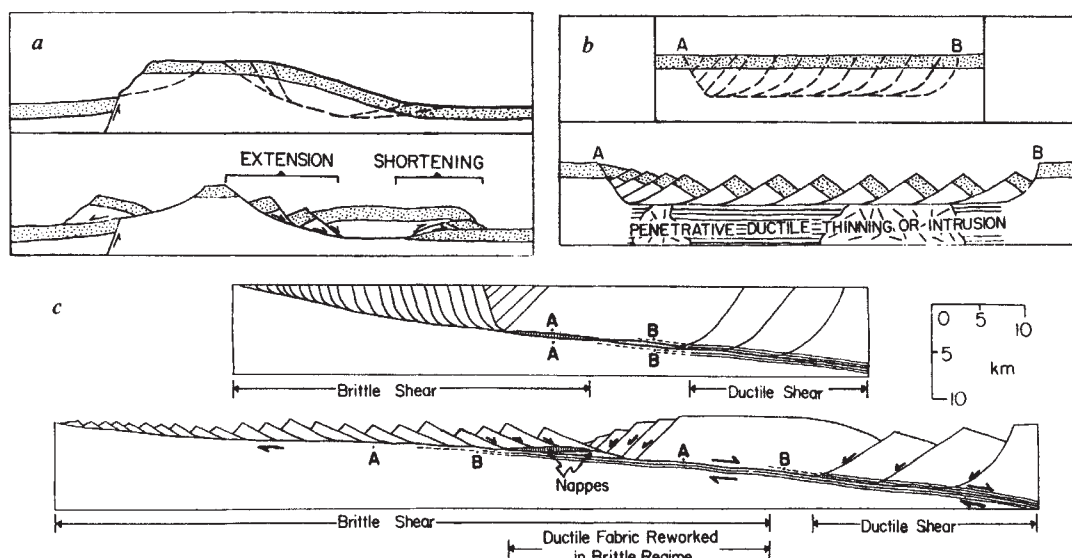
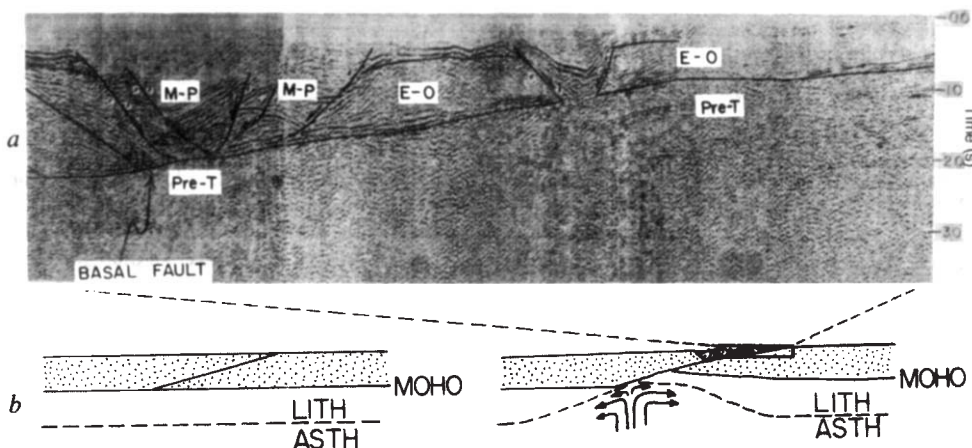


Fig. 4 *a*, North-east trending, 40-km long two-way time section through the Sevier Desert area²⁹. Pre-T, pre-Tertiary; E-O, Eocene-Oligocene; M-P, Miocene-Pliocene. *b*, Lower crust-mantle extrapolation of faults such as that in *a*. Note that the geometry of the Moho is indistinguishable from how it would look if the crust had penetratively stretched. (Seismic line reproduced with permission of the Rocky Mountain Association of Geologists.)



sequence, the fault between the nappes would place younger rocks on top of older rocks, a direct consequence of the fault rooting in the same direction as transport. Had the same juxtaposition occurred along a fault which rooted opposite its transport direction, movement between the nappes would be an expression of true crustal convergence. Large, rooted low-angle normal faults may form zones of ductile shear along deeper segments. If the displacement is large enough, these zones may be shallow and be tectonically reworked in brittle conditions²¹. This scheme of nappe emplacement, combined with superjacent 'domino-style'²⁴ and listric normal faulting, are major mechanisms of extension above the basal fault. Although some aspects of this model are not new^{25,26}, its application to low-angle normal fault terranes in the BRP has been neglected in favour of gravity and *in situ* models.

Distinction between low-angle normal fault terranes and active BRP tectonics has been strongly emphasized by some workers²⁷, and others have related the apparent difference in style (low-angle compared with high-angle) to plate tectonics²⁸. However, the neotectonics of the BRP are apparently sufficiently ill-characterized at depth to entertain the hypothesis that rooted, low-angle normal faults are an important mode of Plio-Quaternary extension, perhaps the fundamental control of its famous topography. For example, the Amargosa Chaos⁴, a striking example of low-angle normal faulting¹, involves mid- to late-Pliocene volcanics, and hence was active well within what some consider to be a province-wide period of high-angle faulting. Further evidence suggesting that the present mechanism may control active BRP tectonics comes from seismic reflection profiles of the Sevier Desert area²⁹. Figure 4*a* and three other profiles reveal a minimum area of ~5,000 km² underlain by a gently west-dipping low-angle normal fault, above which Miocene-Pliocene basin fill and Quaternary basalt flows are offset by normal faults²⁹⁻³¹.

The basal fault in Fig. 4*a* has been interpreted as a reactivated thrust fault²⁹, a common interpretation of low-angle normal faults situated in older thrust terranes^{25,32,33}. However, recent studies of some exhumed thin-skinned extensional areas suggest that pre-existing anisotropies are not a prerequisite for low-angle normal faults⁵, and thrust faults, if present, need not control their geometry. In the Mormon Mountains of southern Nevada (Fig. 1), a Miocene low-angle normal fault event is superposed on a system of Mesozoic thrusts, a situation geometrically identical to that in the Sevier Desert³⁴. The normal faults there truncate the thrusts at high angle, and clearly involve Precambrian crystalline rocks of the thrust autochthon. Based on this study, the normal faults in the Sevier Desert may be unrelated to Mesozoic thrusts prominent in the surface geology of the area.

Figures 3*c* and 4*a* suggest that conservative estimates of province-wide extension (10–30%) based on its proportionality to regional tilts of Tertiary strata may have no meaning³⁵, because large, horizontal translations of rock masses may occur

without appreciable rotation of strata (as well as opposite tilts within the same allochthon). The present model therefore compliments larger estimates (50–100%) deduced by strike-slip fault reconstruction (work in preparation) and crustal thickness arguments²⁰.

Another critical aspect of the model is that it presents an alternative to penetrative ductile thinning or intrusion to accommodate extension at crustal (and mantle) levels not exposed in the BRP. Whereas igneous intrusion may increase the width of the crust, it cannot thin it; at best, it can retain its original thickness, if the intrusions are vertical dykes. Alternatively, it can actually thicken, if the intrusions are sill-like, a property of the large early- to mid-Tertiary batholiths of southern Arizona²². Because intrusion cannot therefore thin deeper crustal levels, it is difficult to envisage it as a major mechanism of BRP extension. Although taffy-like stretching of crustal dimensions does not encounter these geometrical difficulties, there is no compelling evidence to suggest that it is a major process at depth. An attractive alternative, consistent with surface processes, is that large, rooted normal faults extend at a low angle deep into the lithosphere, and that extension at the surface is due to discrete shear between large coherent sheets (Fig. 4*b*). Pull-apart may then ultimately be accommodated by asthenospheric convection rather than any form of stretching.

I thank colleagues at MIT for discussions. This work was supported by NSF grant EAR 7713637 awarded to B. C. Burchfiel.

Received 18 December 1980; accepted 17 March 1981.

1. Wright, L. A. & Troxel, B. W. In *Gravity and Tectonics* (eds de Jong, K. A. & Scholton) 397–407 (Wiley, New York, 1973).
2. Anderson, R. E. *Bull. geol. Soc. Am.* **82**, 43–58 (1971).
3. Misch, P. *Intermountain Ass. petrol. Geol. 11th a. Field Conf. Guidebook*, 17–42 (1960).
4. Noble, L. F. *Bull. geol. Soc. Am.* **52**, 941–1000 (1941).
5. Davis, G. A., Anderson, J. L., Frost, E. G. & Shackelford, T. J. *Geol. Soc. Am. Mem.* **153**, 79–129 (1980).
6. Compton, R. R., Todd, V. R., Zartman, R. E. & Naeser, C. W. *Bull. geol. Soc. Am.* **88**, 1237–1250 (1977).
7. Rehrig, W. A. & Reynolds, S. J. *Geol. Soc. Am. Mem.* **153**, 131–158 (1980).
8. Davis, G. H. *Geol. Soc. Am. Mem.* **153**, 35–78 (1980).
9. Nelson, R. B. *Bull. Am. Ass. petrol. Geol.* **53**, 307–339 (1969).
10. Fritz, W. H. *Nevada Bureau of Mines Map No. 35* (1968).
11. Bick, K. F. *Bull. Utah Geol. miner. Soc.* **77** (1966).
12. Dechert, C. P. thesis, Univ. Washington (1967).
13. Young, J. C. *Intermountain Ass. petrol. Geol. 11th a. Field Conf. Guidebook*, 158–172 (1960).
14. Lee, D. E., Marvin, R. F., Stern, T. W. & Peterman, Z. E. *U.S. geol. Surv. Prof. Pap.* **700-D**, D92–D102 (1970).
15. Armstrong, R. L. *Bull. geol. Soc. Am.* **83**, 1729–1754 (1972).
16. Drewes, H. U. *U.S. geol. Surv. Prof. Pap.* **557** (1967).
17. Moores, E. M., Scott, R. B. & Lumsden, W. W. *Bull. geol. Soc. Am.* **79**, 1703–1726 (1968).
18. Seager, W. R. *Bull. geol. Soc. Am.* **81**, 1517–1538 (1971).
19. Proffett, J. M. Jr. *Bull. geol. Soc. Am.* **88**, 247–266 (1977).
20. Hamilton, W. B. *Pacif. Coast Paleogeogr. Symp.* **2**, 33–70 (1978).
21. Snoke, A. W. *Geol. Soc. Am. Mem.* **153**, 287–334 (1980).
22. Kieth, S. B. *et al. Geol. Soc. Am. Mem.* **153**, 217–268 (1980).
23. Gary, M., McAfee, R. Jr. & Wolf, C. L. (eds) *Glossary of Geology* (American Geological Institute, Falls Church, 1972).
24. Ransome, F. L., Emmons, W. H. & Garrey, G. H. *Bull. U.S. geol. Surv.* **407** (1910).
25. Bally, A. W., Gordy, P. L. & Stewart, G. A. *Bull. Can. petrol. Geol.* **14**, 337–381 (1966).
26. Hunt, C. B. & Mabey, D. R. *U.S. geol. Surv. Prof. Pap.* **494-A** (1966).
27. Reynolds, S. J. & Rehrig, W. A. *Geol. Soc. Am. Mem.* **153**, 159–176 (1980).

28. Thompson, G. A. & Zoback, M. L. *Tectonophysics* **61**, 149–181 (1979).
 29. McDonald, R. E. *Symp. on Cordilleran Hingeline*, 281–317 (Rocky Mountain Association of Petroleum Geologists, 1976).
 30. Hoover, J. D. *Brigham Young Univ. Geol. Stud.* **21**, 3–72 (1974).
 31. Clark, E. *Brigham Young Univ. Geol. Stud.* **24**, 87–114 (1976).
 32. Dahlstrom, C. D. A. *Bull. Can. petrol. Geol.* **18**, 332–406 (1970).
 33. Royse, F. Jr, Warner, M. A. & Reese, D. C. *Deep Drilling Frontiers in Central Rocky Mountains*, 41–54 (Rocky Mountain Association of Petroleum Geologists, 1975).
 34. Wernicke, B. P. *Geol. Soc. Am. Abstr. Prog.* **13**, 113–114 (1981).
 35. Stewart, J. H. *Bull. geol. Soc. Am.* **91**, 460–464 (1980).

Deep seismic soundings in India and the origin of continental crust

K. Roy Chowdhury & R. B. Hargraves*

National Geophysical Research Institute, Hyderabad-500 007, India

The present day continental crust beneath stable cratons is generally thought to be 30–50 km thick^{1,2} although there are divergent views as to its initial thickness and the associated thermal and tectonic regimes^{3,4}. Likewise, the lithosphere is believed to have been thinner in the Archaean, but its relationship to the crust at that time has not been specified. Structural details regarding the Archaean crust revealed by a Deep Seismic Sounding (DSS) study across the South Indian Craton⁵, supplemented by similar data from the Ukrainian Shield¹, show a relationship between crust and mantle which apparently places important constraints on the thermo-tectonic conditions prevailing at the time of formation of the crust. We interpret these details as indicating that at one time the crust constituted the lithosphere, the Moho marking its base.

Figure 1 shows the main geological units of the Indian Peninsular Shield. Apart from narrow coastal strips of younger rocks, the terrain is entirely Precambrian in age. Radiometric data^{6–9} indicate that the development of the South Indian Craton took place 3,400–2,000 Myr ago. Unmetamorphosed dyke-swarms are quite common in the Shield and one such swarm has been dated at $2,420 \pm 250$ Myr (Rb/Sr whole rock isochron)⁸. The Champion Gneiss of Kolar and parts of the Closepet Granites give, however, whole rock and mineral ages around 2,000 Myr (ref. 7), indicating a major, later, thermal event. On the other hand, the rocks of the Cuddapah Basin are reported to be younger than 1,700 Myr (ref. 10). Thus it seems that stable craton was essentially established 2,000 Myr ago.

The resolving power of deep seismic sounding (DSS) developed in the Soviet Union and the more recent deep reflection profiling (DRP) used in the United States is at least an order of magnitude greater than of those based on earthquake or long-range refraction studies, because of the higher frequencies and the correspondingly smaller wavelengths being used to 'sound' the Earth. Whereas seismological investigations use wavelengths ranging from a few hundreds to a few tens of kilometres (free oscillations, surface wave—and body wave—studies), the DSS and the DRP techniques involve shorter waves only a few hundreds of metres long and less. Thus seismic discontinuities can be resolved within the crust to a degree much greater than before. For example, in the DRP technique a vertical resolution of 150 m and a horizontal resolution of 500 m has been reported¹¹.

In DSS^{1,5} mostly wide angle reflections are recorded from discontinuities situated at depths down to and below Moho, the nominal crust–mantle boundary. Suitably selected shot points (typically 40 km apart) and continuous recording up to at least 200-km profile segments (using grouped geophones typically 200 m apart) enable overlapping and reciprocal coverage. The recorded waves include both reflections and refractions from shallow, intermediate and deep boundaries, and diffractions. The wide angle reflections from the Moho show up as high energy arrivals (due to the sudden and large increase in the

reflection coefficient near the wide angle range resulting in almost total reflection) and dominate the records in the distance range of 80–150 km from the shot points.

The DSS data reveal that the crust is broken into blocks by deep-faults. In addition, the structure within the blocks themselves is complex, characterized by a large number of reflector and refractor segments that are up to a few tens of kilometres in length and exhibit irregular dips, even within individual fault-blocks, up to 30°–40°. The Moho segments are more continuous with more regular dips. The simple three-layered crustal model (sediments, granite, gabbro/granulite) inherited from traditional seismology has consequently to be replaced by one that is more complex. The model must contain both vertical and lateral heterogeneities, possibly accompanied by many localized velocity reversals (ref. 1, p. 44, ref. 12).

Figure 2 shows the disposition of the Moho along the 600-km long Kavali–Udipi profile⁵. Here the crust is seen to be cut up into 17 major and minor blocks by deep and intermediate depth faults and thrusts. Many of these faults offset the Moho, up to as much as 7 km (fault 19, between blocks XIII and XIV). The depth to the Moho, in this section, ranges from 34 to 41 km and averages about 38 km. However, within each of these blocks the Moho is essentially horizontal, although short inclined segments are occasionally present.

As seismic profiling yields only apparent dips, inversion of DSS data in terms of the complete crustal structure would require a gridlike ground coverage, not yet available for the region considered here. However, the Kavali–Udipi profile, shown in Fig. 1 was deliberately oriented to be broadly perpendicular to the regional strike. Therefore the dips reported here are believed to be close to their true values. (In Fig. 2 the vertical scale is exaggerated 3:1, thereby accentuating apparent dips).

Figure 3 (no vertical exaggeration) shows the middle part of the Kavali–Udipi crustal section in some detail. The intra-

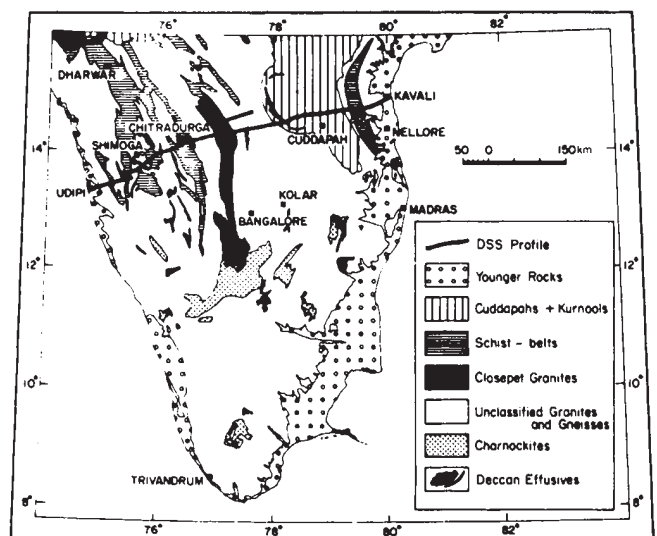


Fig. 1 The Indian Peninsular Region showing the location of the Kavali–Udipi DSS profile (recorded during 1972–75). The Shimoga and the Chitradurga schist-belts in the western part are predominantly sedimentary, rather than of mafic/volcanic type. Radiometric data^{6,7} place them as late Archaean–early Proterozoic and due to their prominent marble–orthoquartzite association these have been termed 'greenstone-like geosynclinal piles'¹⁸. The greenstones proper, dominated by volcanics, are distributed widely across the peninsula, represented by the Nellore schist-belt on the east¹⁸, and numerous remnants (such as Kolar) including a narrow patch on the eastern margin of the Chitradurga belt. The crescent-shaped Cuddapah Basin in the east contains mainly unmetamorphosed and minimally deformed platformal deposits of mid-Proterozoic age¹⁰. On its eastern flank the basin is truncated by a deep thrust⁵ along which the Nellore schist-belt has overridden and overturned the Upper Cuddapahs. Near the western border of the Cuddapah Basin, abundant basaltic dykes and traps are exposed.

*Permanent address: Department of Geological and Geophysical Science, Princeton University, Princeton, New Jersey 08544, USA.

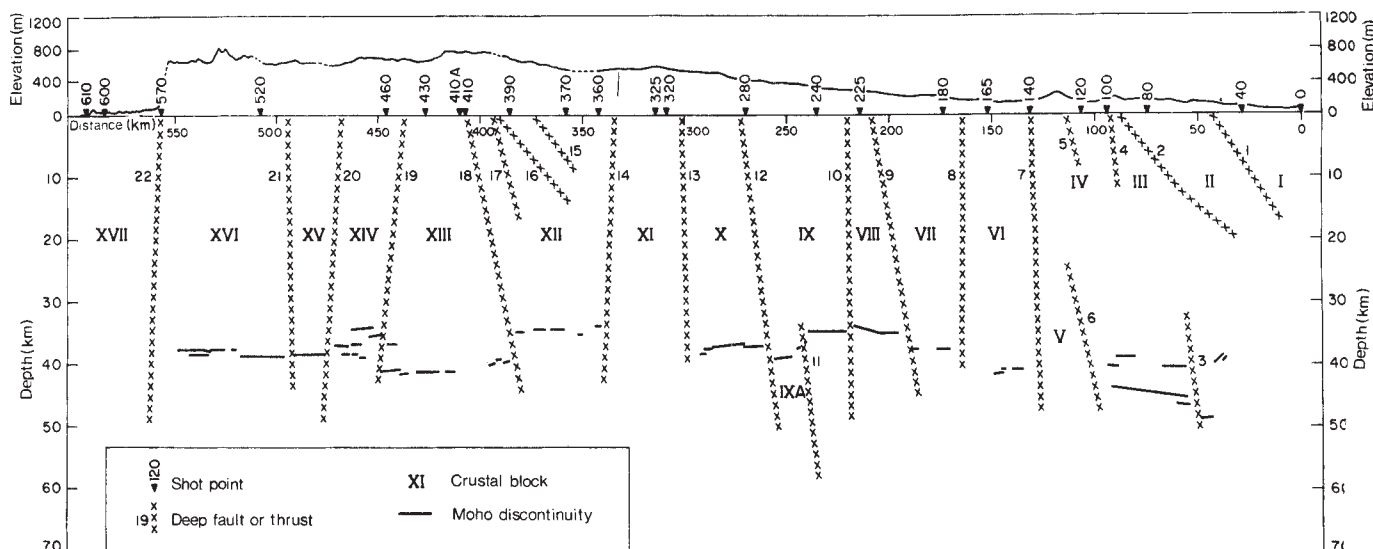


Fig. 2 Crustal thickness and block structure along the Kavali-Udipi DSS profile. As the vertical scale is exaggerated by 3:1 the dips are accentuated.

crustal reflector segments in blocks VI, VIII and X are fairly representative of the rest of the section shown in Fig. 2 and are characterized by rather irregular dips. These may be due to localized folding of crustal layers. This folding, however, has evidently not affected the underlying Moho, which is almost horizontal.

In blocks VII and IX, the fine structure is relatively consistent throughout, with dips up to 20°; yet, here too, the Moho is still essentially horizontal. In these two blocks in particular, the Moho seems to truncate the crustal fabric. Regardless of how this fabric may have originated, the geometry of its structural discordance with respect to the Moho (analogous to an inverted unconformity beneath blocks VII and IX) requires that the Moho, as seen today, formed subsequently.

The above interpretation implies that, at the time of its formation, the Moho constituted a thermo-mechanical boundary between rigid crust and plastic mantle. Any relief on the Moho, such as may have been induced by folding or faulting and tilting of crust, was subsequently erased by ductile flow, whereas the structure was preserved in the cooler crust above.

Ductile flow requires that the then Moho level temperatures were close to the crustal solidus. For anhydrous granulite facies rocks, which are believed to constitute the lower crust, ductile flow might be expected at temperatures around 1,000 °C (cf. 13). Consequently, mean geothermal gradients must have been in the range of 20–30 °C km⁻¹, two to three times those inferred for the present day shields¹⁴. Note that a very similar thermal regime has been inferred from considerations of phase-equilibria in the Lewisian terrain both during the major metamorphism and later at the time of intrusion of Scourie dykes¹³. It is also within the bounds determined from the analysis of *P-T* data for many Archaean high-grade terrains¹⁵.

If the widespread horizontality of the Moho is the result of ductile smearing, then the Moho offsets that are visible today must have arisen later, and their preservation implies that the lithosphere had become cooler by then, with flow possible only at greater depths. This implies that at the time of smearing the Moho was coincident with the base of the lithosphere.

Lithosphere could be described in two ways, depending on the criterion used: 'thermal' lithosphere may be defined as the zone in which heat transfer takes place predominantly by conduction, whereas the thickness of the 'elastic' lithosphere is a function of the change of horizontal shear stress with depth¹⁶. Starting from a fluid sphere, where they have zero thickness, both grow at different rates with progressive cooling and solidification. Although the elastic lithosphere is much thinner beneath the continents today than the thermal lithosphere, in the early Precambrian they could not have differed much. The smearing-out of relief by ductile flow, postulated above, would, strictly speaking, reflect the base of the elastic lithosphere.

Support for our hypothesis is provided by DSS investigations elsewhere. Central and southeastern Europe have been extensively covered by this technique¹. In addition to the traverses across Alpine, Hercynian and Caledonian zones, the areas investigated include the Ukrainian Shield (ref. 1, p 19–23). The latter has been so densely covered by DSS profiles, that a Moho contour map could be compiled (ref. 1, Fig. 25). Figure 4 shows the deep crustal section along part of the profile IP I which extends roughly north-south (right to left) across the Ukrainian Shield. The northern part traverses the exposed part of the Shield, beneath which the Moho is seen to be essentially horizontal, and intra-crustal dips are again suggestive of a structural discordance. To the south, the Moho is seen to be dipping. This area is, however, covered by sediments which thicken towards the metasedimentary complex beneath the Scythian Platform. This complex has been multiply deformed in late Proterozoic and Phanerozoic times (ref. 1, pp. 61, 62). We

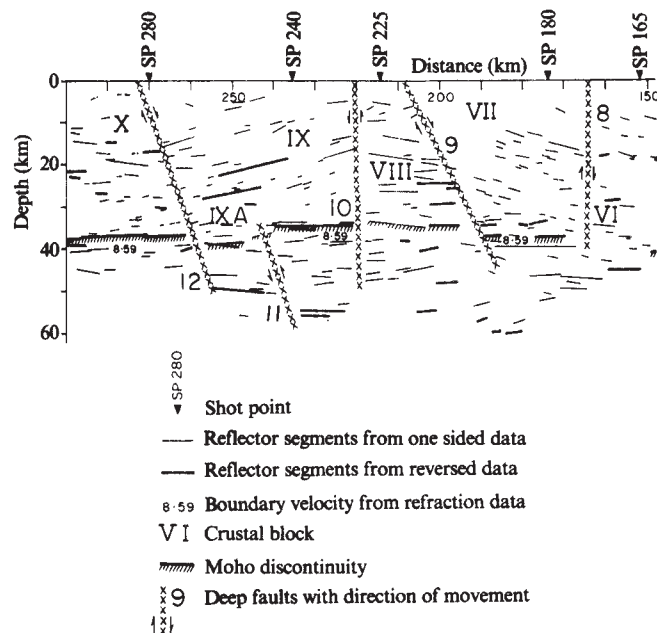


Fig. 3 Detailed crustal structure for the east-central part of the Kavali-Udipi profile (after ref. 5).

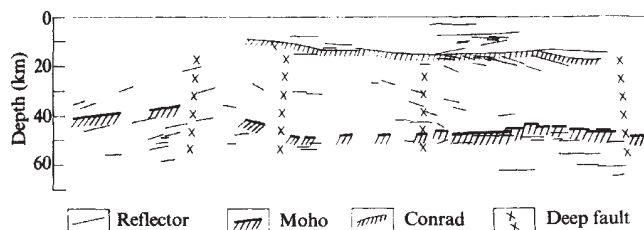


Fig. 4 Simplified crustal section along a part of the profile IP I in the Ukrainian Shield (after ref. 1).

consider that the Moho dips on the left half of Fig. 4 herald the approach to this orogenic zone where deformation occurred when the lithosphere was cooler and thicker, and thus folding of the Moho was preserved.

Preliminary analysis of the near vertical range COCORP data from North Texas, Wyoming and New Mexico in North America has also yielded evidence of a gently dipping Moho¹⁷. The occurrence of this predominantly horizontal structural style has been ascribed to the Moho under the continents being "a buoyancy driven sharp density contrast (re-) established in a region undergoing strong horizontal shear"¹¹. This interpretation is consistent with our hypothesis.

We thank Rishi Narain Singh and Edward S. Grew for helpful discussions. Shankeriah, Jamaluddin and Rana drew the figures. R.B.H. acknowledges the support of an Indo-US Sub-commission fellowship. K.R.C. thanks the director, NGRI for permission to publish.

Received 16 December 1980; accepted 16 April 1981.

1. Jollogub, V. B., Prosen, D. & Militzer, H. (eds) *Geophys. Trans. Hungarian Geophys. Inst.* (spec. edn 1972).
2. Jamford, D. & Prodehl, C. *J. geol. Soc.* **134**, 139–151 (1977).
3. Jondie, K. C. *Plate Tectonics and Crustal Evolution* (Pergamon, New York, 1976).
4. Rye, W. S. *Nature* **249**, 338 (1974).
5. Kaila, K. L. *et al. J. geol. Soc. Ind.* **20**, 307–333 (1979).
6. Crawford, A. R. *J. geol. Soc. Ind.* **10**, 117–166 (1969).
7. Venkatasubramanian, V. S. & Narayana Swami, R. *J. geol. Soc. Ind.* **15**, 77–80, 200–204, 318–319 (1974).
8. Ikramuddin, M. & Stueber, A. M. *Lithos* **9**, 235–240 (1976).
9. Beckinsale, R. D., Drury, S. A. & Holt, R. W. *Nature* **283**, 469–470 (1980).
10. Crawford, A. R. & Compston, W. *J. geol. Soc. Austr.* **19**, 453–464 (1973).
11. Phinney, R. A. *EOS* **59**, 389 (1978).
12. Smithson, S. B. *Contr. Geol. Univ. Wyoming* **17**, 65–75 (1979).
13. O'Hara, M. J. *J. geol. Soc.* **134**, 185–200 (1977).
14. Rao, R. U. M., Rao, G. V. & Narain, H. *Earth planet. Sci. Lett.* **30**, 57–64 (1976).
15. Watson, J. V. *Phil. Trans. R. Soc. A288*, 431–440 (1978).
16. Phillips, R. J. & Ivins, E. R. *Phys. Earth planet. Inter.* **19**, 107–148 (1979).
17. Cook, F. A., Brown, L. D. & Kaufmann, S. *EOS* **59**, 389 (1978).
18. Naqvi, S. M., Rao, V. D. & Narain, H. *Precamb. Res.* **6**, 323–345 (1978).

The earliest known Palaeocene mammal fauna and its implications for the Cretaceous–Tertiary transition

J. David Archibald

Department of Biology and Peabody Museum of Natural History, Yale University, New Haven, Connecticut 06520, USA

Several localities in northeastern Montana, USA, representing the earliest known Palaeocene mammal sites in the world, document a gradual turnover in the mammalian fauna at the Cretaceous–Tertiary boundary. These localities, in combination with previously known latest Cretaceous localities, represent the only well documented terrestrial sequence across the Cretaceous–Tertiary boundary. This sequence argues for gradual terrestrial extinctions and faunal changes at this boundary.

The extensively studied latest Cretaceous Lance local fauna, Lance Formation, Wyoming, USA¹, and Trochu local fauna, Scollard Formation, Alberta, Canada², are dominated by ptilodontoid multituberculates and marsupials. Proteutherian eutherians are somewhat more taxonomically diverse at the Alberta sites than at those in Wyoming, but at neither are eutherians the dominant mammals. These mammalian faunas are considered typically latest Cretaceous. Their composition differs markedly from early Palaeocene (Puercan) faunas.

Localities which yield only typical latest Cretaceous mammals are also known from McCone County, Montana, USA and have been even better sampled from localities immediately to the west in Garfield County³. These localities belong to the Hell Creek faunal-facies which is found throughout most of the vertical and lateral extent of the Hell Creek Formation and which includes sediments deposited in small streams, flood plains and back swamps^{3,4} (Fig. 1).

In marked contrast, a series of three biostratigraphically positioned latest Cretaceous localities (Bug Creek Anthills, Bug Creek West, Harbicht Hill) in McCone County, Montana (Fig. 1) have yielded mammals of a distinctly Palaeocene aspect,

Table 1 Hell's Hollow local fauna, Garfield County, northeastern Montana, USA

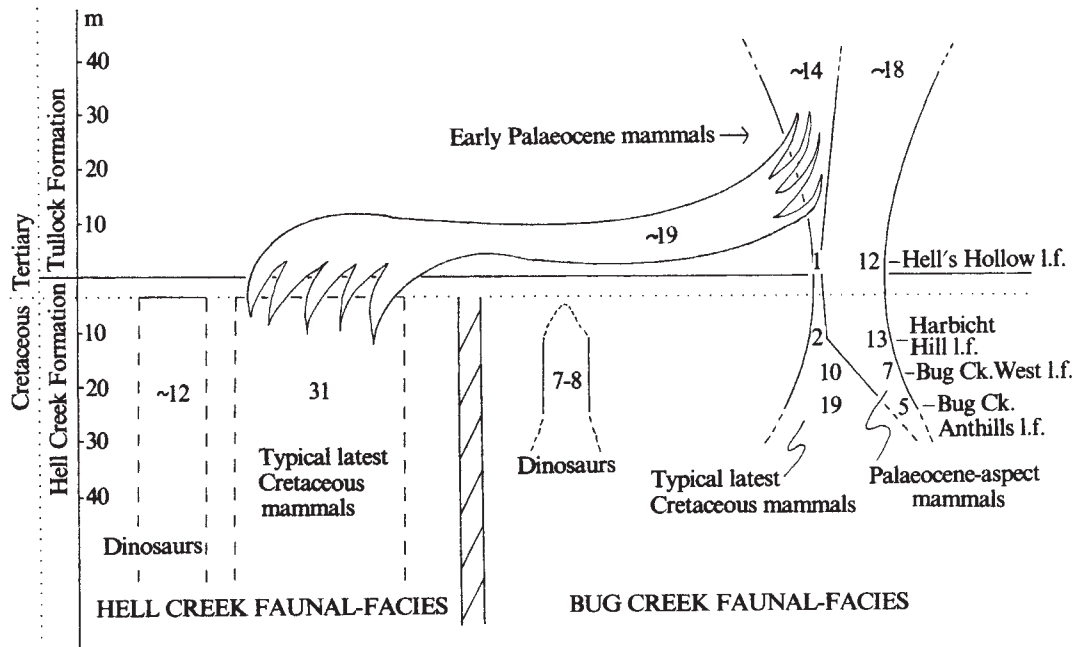
Allotheria—Multituberculata	Theria
Ptilodontoidea	Metatheria—Marsupialia
Neoplagiaulacidae	Didelphidae
<i>Mesodma</i> , n. sp.	<i>Peradectes</i> cf. <i>Peradectes</i>
Taeniolabidoidea	<i>pusillus</i>
Eucosmodontidae	Eutheria—order undetermined
<i>Stygimys</i> aff. <i>Stygimys</i>	(proteutherian)
<i>kuszmauli</i>	Palaeoryctidae
Microcosmodontinae, n. gen.	<i>Procerberus formicarius</i>
et n. sp.	Eutheria—Condylarthra
Taeniolabididae	Arctocyonidae
<i>Caropsalis</i> n. sp.	<i>Protungulatum</i> cf.
Suborder and family undetermined	<i>Protungulatum donnae</i>
<i>Cimexomys minor</i>	<i>Ragnarok harbichti</i> ?
<i>Cimexomys</i> n. sp.	<i>Ragnarok</i> n. sp.
	?Hyopsodontidae
	<i>Oxyprimus erikseni</i>
	Periptychidae
	<i>Mimatuta morgothi</i>

including condylarths, taeniolabidoid multituberculates, a proteutherian ancestral to taeniodonts, and a primate⁵. These localities belong to the Bug Creek faunal-facies which is found only in the upper 24 m of the Hell Creek Formation and which includes sediments deposited in large streams^{3,4} (Fig. 1). Typical latest Cretaceous mammals and dinosaurs occur in this faunal-facies but they become taxonomically less diverse and numerically less abundant up-section as the Palaeocene-aspect mammals increase in both aspects^{3,6}.

The major problem in studying the terrestrial vertebrate transition at the Cretaceous–Tertiary boundary has been the lack of closely spaced localities on both sides of the boundary within a geographically isolated region. The same is true for shallow-water marine macroinvertebrates in North America (K. Waage, personal communication). Early Palaeocene (Puercan) mammal localities were previously discovered in both McCone⁵ and Garfield⁷ counties, but 50 m of section separate these from the latest Cretaceous localities. The discovery of new basal Palaeocene localities in Garfield County separated by only 14 m from latest Cretaceous localities now provides the most closely sampled series of terrestrial vertebrate faunas at the Cretaceous–Tertiary boundary in the world.

The two most fossiliferous of the new basal Palaeocene localities occur in a channel deposit ~5 m above the Cretaceous–Tertiary boundary (as defined by the highest occurrence of dinosaurs) and together comprise the Hell's Hollow local fauna. This local fauna contains 12 or possibly 13 mammalian species and an undetermined number of lower vertebrates. Based on biostratigraphical, and to a lesser extent stratigraphical evidence, this local fauna is older than other early

Fig. 1 Stratigraphical and biostratigraphical sequence in northeastern Montana showing relationships of the Hell's Hollow l.f. (local fauna) to the Bug Creek faunal-facies sequence—Bug Creek Anthills l.f., Bug Creek West l.f., Harbicht Hill l.f. (ref. 5), the Hell Creek faunal-facies and the later early Palaeocene sites. Numbers and widths of bars approximate numbers of species known at various levels. Numbers for Bug Creek faunal-facies dinosaurs are from ref. 6, numbers for the highest early Palaeocene mammals from refs 5 and 9, and the remainder from ref. 3.



Palaeocene faunas known from western North America. Faunally it is most similar to the local fauna which is stratigraphically highest in the Bug Creek faunal-facies—the Harbicht Hill local fauna, which is latest Cretaceous in age (Fig. 1).

Unlike Harbicht Hill, Hell's Hollow lacks dinosaurs. Of the two typical latest Cretaceous mammal species remaining at Harbicht Hill, the multituberculates *Meniscoessus robustus* and *Cimexomys minor*, only the latter species is found in Hell's Hollow. Species of the multituberculates *Mesodma* and *Cimexomys*, the marsupial *Peradectes*, and possibly the proteutherian *Procerberus* that occur in Hell's Hollow may be derived from known typical latest Cretaceous mammals. All except *Peradectes* are represented by the same or closely related species at Harbicht Hill. Hell's Hollow records the earliest definite occurrence of this typically Palaeocene marsupial. The remaining Palaeocene aspect mammals (Table 1) or their ancestors first appear as immigrants in the lowest Bug Creek faunal-facies local fauna (Bug Creek Anthills). Except for the Asian origin of the multituberculate *Catopsalis*⁸, little is known about the origin of these remaining taxa. The species of the multituberculate *Stygimys* and of the condylarths⁹ *Protungulatum*, *Oxyprimus*, *Mimatuta* and questionably *Ragnarok* that are present at Harbicht Hill seem to continue on into Hell's Hollow. *Ragnarok* is also represented by a new species in Hell's Hollow, as is *Catopsalis*. This same species of *Catopsalis* is found at slightly younger localities in Colorado and Wyoming (M. Middleton, personal communication). A new genus and species of microcosmodontine from Hell's Hollow represents the earliest record of this subfamily of eucosmodontid multituberculates.

This earliest known Palaeocene local fauna is clearly a continuation of the latest Cretaceous Bug Creek faunal-facies. In turn, the Hell's Hollow local fauna or similar local faunas gave rise to later faunas of the early Palaeocene in western North America. Although there was taxonomic turnover and possibly a decrease in diversity (as suggested in Fig. 1) during the Cretaceous–Tertiary transition in Montana, the changes were gradual. Although the evidence is not as clear, the Hell Creek faunal-facies may have disappeared rather abruptly as a unit at the close of the Cretaceous. However, about 19 of the 31 species known in this faunal-facies or their close relatives occur at Palaeocene localities. Only the marsupials were decimated, dropping from 13 to 1 species. Such selective extinctions point to relatively normal ecological changes, for catastrophic events would not have been so selective.

These new data, combined with previous information on the changes in the terrestrial biota at the Cretaceous–Tertiary

boundary,⁶ all indicate that this was a non-catastrophic transition. Megaflora, and particularly palynoflora, have been sampled across the Cretaceous–Tertiary boundary at several localities in North America and elsewhere¹⁰. The only consistent pattern of major change that has been recorded is the decimation of the pollen and megaflora within the *Aquilapollenites* Province from western North America and Siberia at the close of the Cretaceous, although a few species belonging to this group continue into the Palaeocene¹⁰. Further, in Garfield County, the palynofloral extinctions occurred ~9 m above the extinction of dinosaurs. Similar incongruities found in Alberta¹¹, south-eastern Wyoming¹² and northern Wyoming (personal communication from P. Gingerich to L. Hickey) also point to a recognizably slow transition of the terrestrial biota at the end of the Cretaceous.

This evidence is at least in part compatible with hypotheses—climatic change, competition with invading mammals of Palaeocene aspect⁶, drops in speciation rates due to habitat redistribution¹³—that argue for a gradual transition. Although it is difficult to estimate the duration of this transition, those that have been suggested place the magnitude in the range of tens or hundreds of thousands of years^{9,11}, not in the tens or hundreds of years suggested by various catastrophic hypotheses^{14,15}. Several of these recent hypotheses have presented data which suggest that an extraterrestrial body struck or closely approached the Earth at the close of the Cretaceous and that its biological effects were catastrophic^{14–16}. The chemical and physical evidence for such an event is intriguing and worthy of further pursuit, but the palaeontological evidence cannot be used to evaluate these hypotheses directly. However, it is only through palaeontological investigation that the biological events at the Cretaceous–Tertiary boundary can be investigated.

Comments and data provided by W. A. Clemens, L. Hickey, J. Ostrom, K. Padian, S. Rachootin, D. Schindel and K. Waage are gratefully acknowledged. Much of the field work was supported by NSF grant DEB 75-21027 to W. A. Clemens.

Received 19 January; accepted 24 April 1981.

1. Clemens, W. A. *Univ. Calif. Publ. geol. Sci.* **48**, 1–105 (1964); **62**, 1–122 (1966); **94**, 1–102 (1973).
2. Lillegraven, J. A. *Paleont. Contr. Univ. Kans.* **50** (Vert. 12), 1–122 (1969).
3. Archibald, J. D. *Univ. Calif. Publ. geol. Sci.* (in the press).
4. Clemens, W. A. & Archibald, J. D. *Mém. Soc. géol. Fr.* **139**, 67–74 (1980); Archibald, J. D. & Clemens, W. A. in *The New Uniformitarianism: Events on a Geologic Scale* (eds Berggren, W. A. & Van Couvering, J.) (Princeton Press, in the press).
5. Sloan, R. E. & Van Valen, L. *Science* **148**, 220–227 (1965); Van Valen, L. & Sloan, R. E. **150**, 743–745 (1965).
6. Van Valen, L. & Sloan, R. E. *Evolut. Theory* **2**, 37–64 (1977).
7. Clemens, W. A. *Science* **184**, 903–905 (1974).
8. Kielan-Jaworowska, Z. & Sloan, R. E. *Acta palaeont. pol.* **25**, 187–197 (1979).
9. Van Valen, L. *Evolut. Theory* **4**, 45–80 (1978).

10. Hickey, L. in *The New Uniformitarianism: Events on a Geologic Scale* (eds Berggren, W. A. & Van Couvering, J.) (Princeton Press, in the press).
11. Lerbekmo, J. F., Evans, M. E. & Baadsgaard, H. *Nature* **279**, 26–30 (1979).
12. Leffingwell, H. A. *Geol. Soc. Am. Spec. Pap.* **127**, 1–64 (1970).
13. Bakker, R. T. in *Patterns of Evolution, as Illustrated by the Fossil Record Vol. 5* (ed. Hallam, A.) 439–468 (Elsevier, Amsterdam, 1977).
14. Alvarez, L. W., Alvarez, W., Asaro, F. & Michel, H. V. *Science* **208**, 1095–1108 (1980).
15. Smit, J. & Hertogen, J. *Nature* **285**, 198–200 (1980).
16. Hsü, K. J. *Nature* **285**, 201–203 (1980).

Paternity and inheritance of wealth

John Hartung*

Department of Anthropology, Harvard University, Cambridge, Massachusetts 02138, USA

One of the oldest conjectures in anthropology is that men transfer wealth to their sister's son when the biological paternity of their 'own' children is in doubt^{1–12}. Because maternity is certain, a man is necessarily related to his sister's son and his brother (see Fig. 1). It is argued here that relatedness to male heirs can be assured by passing wealth to sister's sons or down a line of brothers, whether the prevailing kinship system reckons those brothers matrilineally or patrilineally. It is also argued that when several transfers of wealth are considered, a man's likelihood of being cuckolded need not be unrealistically high¹³ for his successive matrilineal heirs to be more related to him than his successive patrilineal heirs (see Fig. 2). Cross-cultural data on sister's son/brother inheritance¹⁴ and frequency of extramarital sex for females¹⁵ support the hypothesis that men tend to transmit wealth to their sister's son and/or brother when the probability that their putative children are their genetic children is relatively low.

In matrilineal inheritance a man's primary heir is his sister's son. This cross-generation transfer can be direct (25 societies) but more often occurs after wealth has been sequentially inherited by younger brothers (≈ 47 societies), the last of whom leaves it to their mutual sister's son. Although matrilineal inheritance does not prevent a man from distributing significant benefits to his own children^{16,17}, and is thereby not a mirror image of patrilineal inheritance¹⁸ (≈ 500 societies), it is unique in designating major cross-generation heirs related to a male benefactor through an unbroken line of females (counts are of a composite inheritance score derived from ref. 14, as described below).

In 1771, after characterizing traditional societies as patrilineal with regard to inheritance of wealth, status and family membership, John Millar¹ noted an exception: "We read of several nations, among whom it [marriage] is either unknown, or takes place in a very imperfect and limited manner. To a people in this situation it will appear that children have much more connection with their mother than with their father. If a woman has no notion of attachment and fidelity to any particular person, if notwithstanding her occasional intercourse with different individuals... the child which she has born is regarded as a member of her own family... so that if any person was desired to give an account of the family to which he belonged, he was naturally led to recount his maternal genealogy in the female line. [In such nations] females, indeed, are held incapable of enjoying the office of chief, but through them the succession to that dignity is continued; and therefore, upon the death of a chief, he is succeeded not by his own son, but by that of his sister." In 1865 McLennan³ renewed the hypothesis with enthusiasm (indeed, over-enthusiasm): "The connection between these two things—uncertain paternity and kinship through females only, seems so necessary—that of cause and effect—that we may confidently infer the one where we find the other."

In 1961 the first empirical study of matrilineality¹⁹ found that it is distributed independently around the world and is associated with fishing, while patrilineality is associated with pastoralism—an intriguing finding in light of another of Millar's

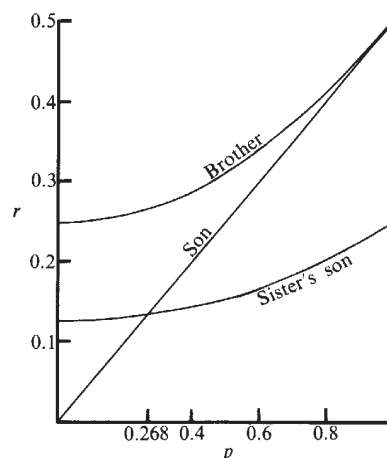


Fig. 1 p = Probability of paternity = the probability that a man's putative offspring by his wife are, in fact, his biological offspring; r = probabilistic degree of relatedness. If p is below 0.268, a man is more related to his sister's son than his 'own' son—see ref. 9 and Fig. 2. A man is more related to his brother whenever $p < 1$. r to son = $\frac{1}{2}p$; r to sister's son equals r through mother to sister to sister's son = $(\frac{1}{2})(\frac{1}{2})(\frac{1}{2})$, plus r through father to sister to sister's son = $(\frac{1}{2}p)(\frac{1}{2}p)(\frac{1}{2})$, which together = $(\frac{1}{2})(\frac{1}{4})(1+p^2)$; r to brother equals r through mother = $\frac{1}{2}$, plus r through father = $\frac{1}{2}p^2$, which together = $\frac{1}{4} + \frac{1}{4}p^2$.

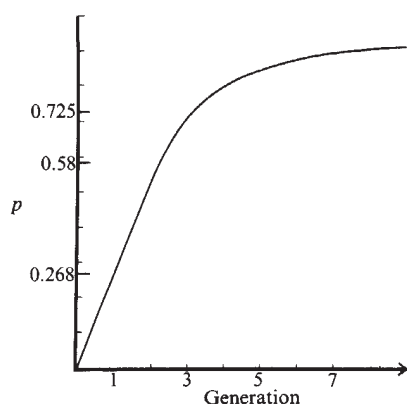
observations¹: "Among the inhabitants of the eastern coast of Tartary it is said, that those tribes which are accustomed to the pasturing of cattle discover some sort of jealousy with regard to the chastity of their women, a circumstance which is looked upon as a matter of perfect indifference by those families in that country who procure their subsistence merely by fishing."

In 1974 Alexander⁷ made the hypothesis explicit with regard to modern evolutionary theory, and in 1978 Greene⁹ calculated the probability of paternity (p) below which husbands must fall if they are to be, on average, more related to their sister's sons than their 'own' sons: $p < 0.268$ (Fig. 1). This threshold is so low because a man's probable relatedness to his sister's son must be reduced by his probable relatedness to his sister through their father, that is, the probability that he is his father's son and she his father's daughter. Unfortunately (for the hypothesis), if a society-wide p were < 0.268 , the average man would have more than twice as much reproductive success by other men's wives as he would by his own wife—a circumstance which, in addition to straining the imagination, could render marriage a maladaptive male strategy (probably better to remain single and expend all of one's efforts on other men's wives).

Fortunately, a less extreme p can fit the paternity hypothesis if more than one generation is considered when contrasting ideal matrilineality against ideal patrilineality (father to son, to son's son, to son's son's son, etc.). The key here is that while a man with a p of 0.8 is probabilistically related to his son by $0.8 \times r$ where r is the degree of relatedness, he is related to his son's son by $0.8^2 \times r$ and to his son's son's son by $0.8^3 \times r$ (where $r = 0.5$, 0.5^2 and 0.5^3 , respectively). That is, the exponent of p increases by a factor of 1 for each transfer to an heir related through a male. In contrast, p affects only the first cross-generation matrilineal transfer. A man's sister's son's sister's son is related to him by half as much as his sister's son, with no additional effect from p . This means, for example, that if p is < 0.725 , a man is more related to his matrilineal heir than his patrilineal heir at three generations (Fig. 2).

Assuming that the paternity hypothesis is viable for reasonable values of p , assessing it would seem a matter of cross-tabulating matrilineal inheritance versus patrilineal inheritance by high versus low extramarital sex for females. However, there is a form of 'delayed' patrilineal inheritance, patrilineal brother inheritance, which can be seen as an alternative adaptation to low p . Given any $p < 1$, a man is more related to his brother than to any other potential male heir (Fig. 1). Because p cannot be 1, one might expect all patrilineal benefactors to follow the pattern

* Present address: 2G, 160 Claremont Ave, New York, New York 10027, USA.



Generation	r to sister's son _i	r to son _i	Threshold at $p =$
1	$(\frac{1}{2})(\frac{1}{4})(1+p^2)$	$\frac{p}{2}$	0.268
2	$(\frac{1}{2})^2(\frac{1}{4})(1+p^2)$	$(\frac{p}{2})^2$	0.58
3	$(\frac{1}{2})^3(\frac{1}{4})(1+p^2)$	$(\frac{p}{2})^3$	0.725
i	$(\frac{1}{2})^i(\frac{1}{4})(1+p^2)$	$(\frac{p}{2})^i$	$4p^i - p^2 - 1 = 0$

Fig. 2 Solving for p given i , $4p^i - p^2 - 1 = 0$, where $i = \text{ith}$ generation, derives the probability of paternity below which a man is more related to his cross-generation matrilineal heir (sister's son_i) than his cross-generation patrilineal heir (son_i) at successive generations. For example, if p is < 0.725 , a man is likely to be more related to his third generation matrilineal heir than to his son's son. (I thank P. Greene for deriving this formula.)

of giving to their younger brother(s) with the stipulation that the wealth subsequently return to their son(s). However, at an average father's age of death, his son is near his reproductive prime while his brother is past that stage. Accordingly, a son's higher reproductive value²⁰ will compensate, in terms of effect on benefactor's inclusive fitness²¹, when p is < 1 but not $\ll 1$.

Why adaptation to low p should be patrilineally reckoned brother inheritance in some societies (≈ 70 cases) and matrilineal inheritance in other societies (≈ 72 cases) is unclear. The absolute relatedness of a brother is higher than that of a sister's son, but in patrilineal brother inheritance wealth eventually reverts to a man's own child. There seems to be a trade-off between absolute relatedness to heir and the number of transfers for which relatedness is assured (with matrilineal inheritance assuring some relatedness indefinitely). Perhaps the two strategies reflect a difference in p , or economic differences which affect the duration of effectiveness of inherited wealth, or perhaps patrilineal brother inheritance is simply as 'matrilineal' as an otherwise patrilineal society can get. Whatever the reason, given brother inheritance, whether in a matrilineal or patrilineal context, any inclusive fitness enhancing effect of wealth^{8,22-24} resides for the longest possible time amongst most highly related males (for emphasis on males see refs 8, 22, 25, 26). The key here is that, for present purposes, patrilineal brother inheritance has more in common with matrilineal inheritance than with patrilineal inheritance. Thus, one should group patrilineal brother inheritance with matrilineal inheritance when testing the hypothesis that men stipulate an heir whose relatedness is less affected by low p in societies where p is relatively low.

Murdock and White have selected 'The Standard Cross-Cultural Sample'²⁷ by taking one society from each of the world's 186 major culture areas, giving a sample which maximizes independence and minimizes effects from cultural diffusion. Separate codes for inheritance of real property (land,

house) and movable property (such as livestock, money) are as follows^{14,28,29}:

- 0: No information.
- 1: No land rights/no inherited movable property/no rule governing the transmission of same.
- 2: Matrilineal inheritance by a sister's son or sons.
- 3: Matrilineal inheritance by heirs who take precedence over sisters' sons (such as younger brother).
- 4: Inheritance by children, but with daughters receiving less than sons.
- 5: Inheritance by children of either sex or both.
- 6: Inheritance by patrilineal heirs who take precedence over sons (such as younger brother).
- 7: Patrilineal inheritance by son or sons.

For present purposes code 3 is functionally the same as code 2—though brother is not stipulated as the only possibility, he is the most likely heir^{16,18}. Where this is not the case, the pattern still qualifies as having heirs unaffected by low p , because any matrilineal heir must be related through the mother and because

PRIMARY HEIR(S)	
Brother/sister's son	Own children
A 6 Nyakyusa Nkundo Lesu Trobrianders Saramacca Crow	B 5 Romans Vietnamese Kwoma Trukese Aymara
C 1 Siuai	D 10 Kikuyu Egyptians Turks Irish Lapps Lepcha Lakher Vedda Javanese Iban

Fig. 3 Inheritance by probability of paternity. $N = 22$; Fisher exact probability = 0.0298; $\gamma = 0.846$ (considering the two inheritance forms as 'high' and 'low' with regard to effect of p on relatedness to heir); $\phi = 0.488$. The Huichol (see text) are rated 'high' on female extramarital sex and would go in cell A or B if included—either way the figure would retain significance below 0.05. This significance depends on grouping patrilineal brother inheritance with matrilineal inheritance. If the contrast were matrilineal versus own children/patrilineal brother, the Nyakyusa, Nkundo and Huichol would fall in cell B; giving cell A = 4, cell B = 8, cell C = 1, cell D = 10: Fisher exact probability = 0.162. If the contrast were simply matrilineal versus own children, the Nyakyusa, Nkundo and Huichol would drop out, giving cell A = 4, cell B = 5, cell C = 1, cell D = 10: Fisher exact probability = 0.0893. Murdock expressed doubt about his inheritance codes¹⁴. In response to an inquiry he wrote that he did not have reservations about the accuracy of the data, but felt that the codes were 'too crude' (personal communication). Future research would profit from refining the inheritance codes. For example, a man could extend considerable assistance to his brother, both during his lifetime and through inheritance and still follow the norms of a society coded in inheritance category 'own children'. According to the paternity hypothesis as presented here, societies in cell B would be expected to follow that pattern more than societies in cell D, and some cell B societies may have sufficient 'brother investment' to qualify for cell A.

wealth eventually goes to sister's son. Societies with code 6, however, must be checked to see whether brother is the heir. Of the 23 cases with usable codes for inheritance (see below) and frequency of female extramarital sex, three contain a 6 in their composite inheritance score: the Nyakyusa (6 and 6), the Nkundo (6 and 6) and the Huichol (5 and 6). These were checked in the Human Relations Area Files²⁸ to see whether brother is the preferred heir. The major ethnographer for the Nyakyusa puts the case succinctly³⁰: "A man's heir is his full brother next to him in age, and only when all the full brothers of a family have died does the property pass to the eldest son of the eldest brother." For the Nkundo, cases were reported of brother taking precedence over son, of sister's son taking precedence over son, and of brother taking precedence over sister's son³¹. As put by Murdock³², the Nkundo "lie wedged between the matrilineal Central Bantu in the south and the patrilineal Equatorial Bantu in the north. Being unusually well described, they shed considerable illumination on the process whereby the transition from the one to the other form of social organization has occurred in this general region." Perhaps the Nkundo would have to be coded 'none of the above', but it is clear that their system provides substantial brother, if not matrilineal, inheritance. The case for the Huichol is not as clear. Movable property goes to children, but real property goes communally to a group of family members that probably, but not necessarily, contains brother(s)³³. Accordingly, the Huichol were left out of the analysis.

To maximize sample size and variable dichotomy, the contrast for inheritance norms was made between: (1) societies that stipulate heirs necessarily related to male benefactors through females (heirs whose *r* to benefactor is relatively unaffected by low *p*), and (2) societies that exclude such heirs and stipulate only those heirs whose *r* to benefactor is highly affected by low *p*. Accordingly, societies were grouped into: (1) those with brother/matrilineal inheritance for either type of property (code 2, 3 or 6, where 6 has been checked—see above), or (2) those for which benefactor's own children are the only stipulated heirs for both types of property (code 4, 5 or 7).

Broude and Greene¹⁵ have coded frequency of extramarital sex for females as: (1) universal, moderate, and (2) rare, not common³⁴. As would be predicted by the paternity hypothesis, the appropriate cross-tabulation (Fig. 3) indicates an association between relatively frequent female extramarital sex and a cultural norm that allows men to designate heirs whose relatedness is relatively unaffected by low probability of paternity.

It is assumed that the behaviours relevant to this analysis have little heritability across same-sex individuals within a society, and negligible heritability by sex across societies—that is, it is presumed that differences in individual behaviour are phenotypic responses to different conditions, that the behaviours in question have been too important for too long for there to be much genetic variance left, and that extant within-sex genetic variance is an irrelevant variable (compare refs 26, 35 with 36).

I thank D. Aberle, R. D. Alexander, N. Chagnon, I. DeVore, M. Dickeman, P. Ellison, P. Greene, W. Irons, P. Harvey, G. P. Murdock, J. Shepherd, B. Smuts, P. van den Berghe and J. Whiting for their thoughts and assistance. Particular thanks are due to N. Schmidt for her impetus in re-acquiring the HRAF at Harvard.

Received 2 September 1980; accepted 8 April 1981.

1. Millar, J. *Observations Concerning the Distinction of Ranks in Society*, 30, 33, 42 (Murray, London, 1771).
2. Bachoffen, J. J. *Das Mutterrecht* (Benno Schwabe, Basel, 1861).
3. McLennan, J. F. *Primitive Marriage*, 161 (Black, Edinburgh, 1865).
4. Morgan, L. H. *Ancient Society* (Holt, New York, 1877).
5. Junod, H. *Life of a South African Tribe* (Attinger Feres, Neuchâtel, 1913).
6. Kahn, M. C. *Djuka: the Bush Negroes of Dutch Guiana* (Viking, New York, 1931).
7. Alexander, R. D. *A. Rev. Ecol. Syst.* 5, 324–383 (1974).
8. Hartung, J. *Curr. Anthropol.* 17, 4, 607–622 (1976).
9. Greene, P. J. *Am. Ethnol.* 5, 1, 151–159 (1978).
10. Kurland, J. A. in *Evolutionary Biology and Human Social Behavior* (eds Chagnon, N. A. & Irons, W.) (Duxbury, North Scituate, Massachusetts, 1979).
11. Flinn, M. in *Natural Selection and Social Behavior: Recent Research and New Theory* (eds Alexander, R. D. & Tinkle, D. W.) (Chiron, New York, 1981).
12. Gaulin, J. C. & Schlegel, A. *Ethol. Sociobiol.* 1, 301–309 (1980).
13. Swetnam, J. J. & Knack, M. C. *Am. Anthropol.* 81, 108 (1979).

14. Murdock, G. P. *The Ethnographic Atlas* 1, 170 case punch card version with codebook, HRAF (University of Pittsburgh Press, 1967).
15. Broude, G. J. & Greene, S. J. *Ethnology* 15, 409–429 (1976).
16. Richards, I. A. in *African Systems of Kinship and Marriage* (eds Radcliffe-Brown, A. R. & Forde, D.) 207–251 (Oxford University Press, London, 1950).
17. Panoff, M. *Ethnology* 15, 175–188 (1976).
18. Schneider, D. & Gough, K. (eds) *Matrilineal Kinship* (University of California Press, 1961).
19. Aberle, D. in *Matrilineal Kinship* (eds Schneider, D. & Gough, K.) 655–727 (University of California Press, 1961).
20. Fisher, R. A. *The Genetical Theory of Natural Selection* (Oxford University Press, 1930).
21. Hamilton, W. D. *J. theor. Biol.* 7, 1–52 (1964).
22. Irons, W. in *Evolutionary Biology and Human Social Behavior* (eds Chagnon, N. A. & Irons, W.) (Duxbury, North Scituate, Massachusetts, 1979).
23. Keyfitz, N. *Applied Mathematical Demography* (Wiley, New York, 1979).
24. Frisch, R. *Science* 199, 22–30 (1978).
25. Hartung, J. *Curr. Anthropol.* (in the press).
26. Hartung, J. in *Natural Selection and Social Behavior: Recent Research and New Theory* (eds Alexander, R. D. & Tinkle, D. W.) (Chiron, New York, 1981).
27. Murdock, G. P. & White, D. R. *Ethnology* 8, 329–369 (1969).
28. HRAF, Atlas codebook (see ref. 14).
29. Textor, R. B. *A Cross-Cultural Summary*, 112 (HRAF, New Haven, 1966).
30. Wilson, M. *Good Company: a Study of Nyakyusa Age Villages*, 35 (Oxford University Press, London, 1951).
31. Hulstaert, G. E. *Le Mariage des Nkundo* (G. van Campenhout, Bruxelles, 1928).
32. Murdock, G. P. *Africa—Its People and their Culture History*, 284–289 (McGraw-Hill, New York, 1959).
33. Zingg, R. M. *The Huichols: Primitive Artists* (Stechert, New York, 1938).
34. Broude, G. J. *Behav. Sci. Res.* 4, 227–262 (1976).
35. Hartung, J. *Curr. Anthropol.* 21, 1, 131 (1980).
36. Lumsden, C. J. & Wilson, E. O. *Genes, Mind and Culture: The Coevolutionary Process* (Belknap/Harvard Press, Cambridge, 1981).

Sustained potentiation of transmitter release by adrenaline and dibutylrly cyclic AMP in sympathetic ganglia

K. Kuba*, E. Kato†, E. Kumamoto*, K. Koketsu† & K. Hirai†

* Department of Physiology, Saga Medical School, Nabeshima, Saga 840-01, Japan

† Department of Physiology, Kurume University School of Medicine, Kurume 830, Japan

Cyclic AMP¹ is known to mediate actions of various hormones and neurotransmitters in a variety of cells^{2–5}. Its physiological role, however, in the transmitter release mechanism of vertebrate synapses is still controversial^{6–9}. We report here evidence for the sustained, acceleratory effects of adrenaline and dibutylrly cyclic AMP on transmitter release in bullfrog sympathetic ganglia. The amplitude and quantal content of the fast excitatory postsynaptic potential (e.p.s.p.) and the frequency of miniature excitatory postsynaptic potential (m.e.p.s.p.) increased markedly for a long time after pretreatment with adrenaline or dibutylrly cyclic AMP. These results suggest a significant modulatory role for adrenaline, presumably linked with a cyclic AMP system, in the transmitter release mechanism, which may serve as a mechanism for neuronal plasticity in the peripheral autonomic nervous system.

The ninth or tenth paravertebral sympathetic ganglia of the bullfrog (*Rana catesbeiana*) were isolated. B-type neurones were studied using a conventional intracellular recording technique¹⁰. Microelectrodes were filled with 3 M KCl or 1 M K₃-citrate (tip resistance 25–100 MΩ). The fast e.p.s.p. generated by the nicotinic action of acetylcholine (see ref. 11) was recorded in a low Ca²⁺-high Mg²⁺ solution (Ca²⁺ 0.7–0.9 mM, Mg²⁺ 5.4–6.5 mM). The quantal content of the fast e.p.s.p. was calculated by both the variance and failure methods¹² and the average taken; m.e.p.s.ps were recorded in a high-K⁺ (10 mM) solution. The composition of normal Ringer was described previously¹⁰ and experiments were done at 20–24 °C.

Figure 1 shows the effects of adrenaline on the fast e.p.s.p. Perfusion of the sympathetic ganglion with a solution containing adrenaline (100 μM) decreased the amplitude of the fast e.p.s.p. (65 ± 7% (s.e.) of the control, *n* = 5) and its quantal content

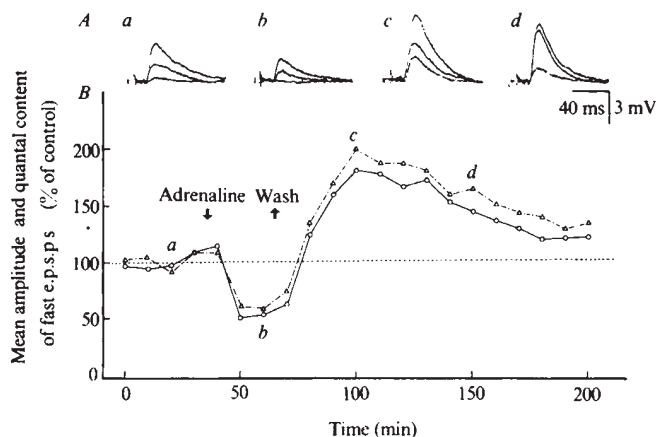


Fig. 1 *A*, effects of adrenaline (100 μ M) on the fast e.p.s.p.s. recorded in a low Ca^{2+} -high Mg^{2+} solution. The amplitude of fast e.p.s.p.s. in each condition fluctuated in a quantal manner for their low quantal contents. *B*, time course of the adrenaline action. Each point is the mean amplitude ($\circ$) or quantal content (Δ) calculated from 200 fast e.p.s.p.s. Quantal contents before application of adrenaline ranged from 1.6 to 2.0. Downward and upward arrows indicate the beginning and end, respectively, of the application of adrenaline. The duration of the 'after-potential' of the fast e.p.s.p. by adrenaline observed in this cell was apparently shorter than that induced by dibutyryl cyclic AMP in the other cell (shown in Fig. 2). However, this difference seems to be mainly due to the variability of cell responses, and not to the specificity of the stimulating agents. In fact, a much longer potentiation (~ 5 h) by adrenaline was seen in other cells (see text). The points *a*, *b*, *c* and *d* in *B* correspond to *a*, *b*, *c* and *d* in *A*, respectively.

(74 $\pm$ 7% of the control, $n = 5$). Ten to fifteen minutes after washing with an adrenaline-free solution, the amplitude and quantal content of the fast e.p.s.p. recovered and rose to 155% of the control ($\pm 8\%$, $n = 5$, the value at 30 min after wash) and to 170% ($\pm 14\%$, $n = 5$), respectively. This potentiation of the fast e.p.s.p. continued for $>2-5$ h (Fig. 1*B*) and was observed with an adrenaline concentration as low as 1 μ M. Similar modes of adrenaline action on the fast e.p.s.p. were also observed in Ringer (at normal Ca^{2+} concentration) containing D-tubocurarine (28–43 μ M). The quantal size of the fast e.p.s.p., the resting potential and membrane resistance remained almost unchanged throughout the adrenaline actions. These results, together with the lack of effects of adrenaline on acetylcholine-induced depolarization of the ganglion cell membrane, indicate the presynaptic bimodal actions of adrenaline which are quite different from the predominantly depressant effects of adrenaline in mammalian sympathetic ganglia¹³. The delayed onset of the facilitatory effect of adrenaline can be accounted for by the existence of the strong inhibitory action of adrenaline during exposure to it. Thus, without the facilitatory action, the magnitude of the inhibition in the presence of adrenaline would have been much more intense than that observed.

The inhibitory action of adrenaline (10 μ M) tended to be antagonized by α -antagonists of the adrenoreceptor, such as phentolamine (10 μ M) and phenoxybenzamine (10 μ M), but not by a β -antagonist, propranolol (10 μ M). On the other hand, the 'after-potential' of the fast e.p.s.p. by adrenaline was relatively small or not seen in the presence of either α - or β -antagonist. Furthermore, both dopamine (known to be an α -agonist; 10 μ M) and isoprenaline (β -agonist; 10 μ M) significantly potentiated the fast e.p.s.p. Thus, it seems premature to determine the type of adrenoreceptor responsible for the facilitation of transmitter release until quantitative and kinetic analyses are made. Alternatively, the receptor for the facilitation mechanism might be a 'primitive' type which cannot be classified into α - or β -type.

The fast e.p.s.p. amplitude tended to be slightly reduced during treatment with dibutyryl cyclic AMP ($N^6,N^{2'}\text{-O}$ -dibutyryl adenosine 3', 5'-monophosphate; 0.8–1 mM), which is

known to be more permeable to the cell membrane than cyclic AMP (Fig. 2*Ab*). Five to fifteen minutes after the removal of dibutyryl cyclic AMP, the fast e.p.s.p. amplitude increased gradually to 165% of the control ($\pm 15\%$, $n = 5$, the value at 30 min after wash). This was accompanied by an increased quantal content (165 $\pm$ 13% of the control, $n = 5$) without a significant change in the quantal size, resting membrane potential or resistance. Furthermore, 'after-potential' was seen at a concentration as low as 0.1 mM. Thus, dibutyryl cyclic AMP has a presynaptic action. As found in the 'after-effect' of adrenaline, the potentiation by dibutyryl cyclic AMP persisted for $>2-5$ h (Fig. 2*c, d*). Similar effects of dibutyryl cyclic AMP were also seen with the fast e.p.s.p. recorded from a curarized ganglion cell in Ringer.

The presynaptic facilitation by dibutyryl cyclic AMP (0.8–1.0 mM) was not observed during treatment for up to 30 min with the drug. By contrast, pretreatment with dibutyryl cyclic AMP even for 5–10 min caused the 'after-potential' of transmitter release. Thus, the presence of dibutyryl cyclic AMP at the outer surface of the terminal membrane seems to inhibit the facilitatory mechanism or its initiation. Cyclic AMP, being less permeable to the cell membrane, AMP, ATP and adenosine (4 mM) did not produce 'after-potential' of the fast e.p.s.p. Furthermore, 3-isobutyl-1-methylxanthine (IBMX, 20 μ M; known to be a phosphodiesterase inhibitor) increased the fast e.p.s.p. amplitude in the ganglion cell which was pretreated with adrenaline (10 μ M) 30–40 min before IBMX application.

To test whether the spontaneous release of transmitter is enhanced by adrenaline and dibutyryl cyclic AMP, as with evoked release, the effects of these agents on the m.e.p.s.p. frequency were observed. Although a change in the m.e.p.s.p. frequency in the presence of adrenaline (100 μ M) was variable among cells, it was clearly enhanced in many cells after the removal of adrenaline (176 $\pm$ 34% of the control at 15 min, $n = 8$). Similarly, the m.e.p.s.p. frequency increased after the application of dibutyryl cyclic AMP (1–4 mM) and its subsequent removal (172 $\pm$ 27%, $n = 9$). These potentiations of spontaneous transmitter release continued for many hours.

Our experimental findings provide evidence of the two hypothetical, but important concepts for the mechanism of vertebrate synaptic transmission. First, the similar modes of presynaptic facilitation by adrenaline and dibutyryl cyclic AMP, but not by less permeable cyclic AMP and other nucleotides, suggest that the action of adrenaline is mediated by endogenous cyclic AMP

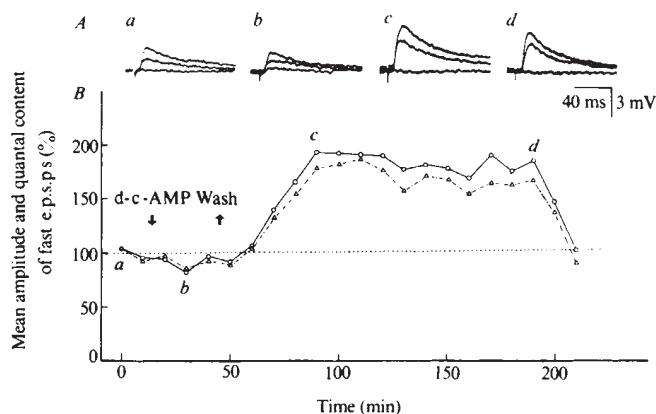


Fig. 2 *A*, effects of dibutyryl cyclic AMP (d-c-AMP) on the fast e.p.s.p.s. recorded in a low Ca^{2+} -high Mg^{2+} solution. Fluctuations of the fast e.p.s.p. amplitude were seen in each condition for their low quantal content. *B*, time course of the dibutyryl cyclic AMP action. The mean amplitude and quantal content of the fast e.p.s.p.s. before application of the drug were taken to be 100% in the ordinate. Control quantal contents ranged from 0.7 to 1.0. Each point is the mean of the amplitude ($\circ$) or quantal content (Δ) calculated from 200 fast e.p.s.p.s. Downward and upward arrows indicate the beginning and end of the treatment with dibutyryl cyclic AMP (1 mM), respectively.

and a related metabolic process. It has been shown that catecholamines increase cyclic AMP in bovine¹⁴ and rat^{15,16} whole sympathetic ganglia, although such an increase in the preganglionic terminals is not yet proved. The latter possibility might be substantiated by the recent observation that dopamine raised the cyclic AMP level in the synaptosome of the guinea pig brain¹⁷. Thus, adrenaline may raise the cyclic AMP level in the preganglionic terminals, thereby triggering an unidentified mechanism that would eventually modify the Ca^{2+} transport mechanism at the presynaptic terminal membrane^{18,19}, intracellular Ca^{2+} metabolism²⁰ or transmitter synthesis². The presynaptic facilitation by serotonin and cyclic AMP in *Aplysia* ganglia was explained by an increase in the active Ca^{2+} conductance during an action potential^{18,19}. However, the mechanism of the potentiation seen in the sympathetic ganglia seems to involve at least either an increased level of the intracellular free Ca^{2+} or a raised amount of releasable transmitter, as both the spontaneous and evoked releases of transmitter were enhanced.

Received 3 December 1980; accepted 8 April 1981.

1. Sutherland, E. W. & Rall, T. W. *J. biol. Chem.* **232**, 1065–1076 (1958).
2. Greengard, P. *Nature* **260**, 101–108 (1976).
3. Nathanson, J. A. *Physiol. Rev.* **57**, 157–256 (1977).
4. Rasmussen, H. & Goodman, D. B. P. *Physiol. Rev.* **57**, 421–509 (1977).
5. Phillips, J. W. *J. Can. Sci. Neurobiol.* **4**, 151–195 (1977).
6. Ginsborg, B. L. & Hirst, G. D. S. *J. Physiol., Lond.* **224**, 629–645 (1972).
7. Miyamoto, M. D. & Breckenridge, B. M. *J. gen. Physiol.* **63**, 609–624 (1974).
8. Standaert, F. G., Dretchen, K. L., Skirbol, L. R. & Morgengroth, V. H. III. *J. Pharmac. exp. Ther.* **199**, 553–564 (1976).
9. Kupfermann, I. A. *Rev. Physiol.* **42**, 629–641 (1980).
10. Kato, E. & Kuba, K. *J. Physiol., Lond.* **298**, 271–283 (1980).

The other significant finding is the prolonged maintenance of the presynaptic facilitation, suggesting a plasticity in the peripheral autonomic nervous system. When impulses from the central nervous system activate the postganglionic neurone soma and/or the internuncial adrenergic cells (small intensely fluorescent cells), these would liberate catecholamines. An increased level of catecholamines around the preganglionic terminals might set off the facilitatory mechanism of transmitter release process, and result in the enhanced efficacies of not only the fast (nicotinic), but also the slow (muscarinic) synaptic transmission¹¹ for many hours. The presynaptic facilitation by adrenaline and dibutyl cyclic AMP in the bullfrog sympathetic ganglia is quite different from the postsynaptic mechanism for the potentiation of the slow e.p.s.p. by dopamine and cyclic AMP in mammalian ganglia^{21,22}, possibly due to species difference. As already discussed, however, the slow e.p.s.p. would also be augmented if the transmitter release is potentiated.

11. Kuba, K. & Koketsu, K. *Prog. Neurobiol.* **11**, 77–169 (1978).
12. Del Castillo, J. & Katz, B. *J. Physiol., Lond.* **124**, 560–573 (1954).
13. Christ, D. D. & Nishi, S. *J. Physiol., Lond.* **213**, 107–117 (1971).
14. Kebabian, J. W. & Greengard, P. *Science* **174**, 1346–1349 (1971).
15. Cramer, H., Johnson, D. G., Hanbauer, I., Silberstein, S. D. & Kopin, I. J. *Brain Res.* **53**, 97–105 (1973).
16. Brown, D. A., Caulfield, M. P. & Kirby, P. J. *J. Physiol., Lond.* **290**, 441–451 (1979).
17. Kobayashi, K., Kuroda, Y. & Yoshioka, M. *J. Neurochem.* **36**, 86–91 (1981).
18. Shimahara, T. & Tauc, L. *J. Physiol., Paris* **74**, 515–519 (1978).
19. Klein, M. & Kandel, E. R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3512–3516 (1978).
20. Tada, M., Yamamoto, T. & Tonomura, Y. *Physiol. Rev.* **58**, 1–79 (1978).
21. Libet, B., Kobayashi, H. & Tanaka, T. *Nature* **258**, 155–157 (1975).
22. Kobayashi, H., Hashiguchi, T. & Ushiyama, N. S. *Nature* **271**, 268–270 (1978).

Functional EGF receptors are present on mouse embryo tissues

Eileen D. Adamson*, M. Jane Deller† & Joseph B. Warshaw‡

*La Jolla Cancer Research Foundation, 2945 Science Park Road, La Jolla, California 92037, USA

†Department of Zoology, South Parks Road, Oxford University, Oxford OX1 3PS, UK

‡Division of Perinatal Medicine, Department of Pediatrics and Obstetrics and Gynecology, Yale University School of Medicine, New Haven, Connecticut 06510, USA

Epidermal growth factor (EGF) is a small polypeptide (6,045 molecular weight) well recognized as a mitogen affecting a variety of cell types (see refs 1–3 for reviews). Its mitogenic effect is mediated through specific high-affinity receptors on the plasma membrane, and after ligand binding, some (but not all) cells respond in a variety of ways which lead to either increased growth rate or differentiation. Although the initial description of EGF by Cohen⁴ demonstrated a maturational influence on eye-opening and tooth eruption in newborn mice, there has been little investigation of the effect of EGF on developing fetal tissues. Nexø *et al.*⁵ have provided evidence that EGF and its receptors are present in mouse embryos as early as day 13 of gestation, especially in the secondary palate where EGF may have a role in palate closure, and EGF receptors have also been described in the human placenta⁶ and in fetal rabbit lung⁷. Thus, the limited data available suggest that EGF may play a part in fetal growth and development of specialized function. Here, we describe the binding and mitogenic activity of EGF both *in vitro* and *in vivo* in several fetal mouse tissues.

Tissues dissected from embryos at days 11–18 of gestation (day 1 being the day that a copulation plug is found) were subjected to ¹²⁵I-EGF *in vitro* binding studies (Table 1). For most tissues the time of maximum concentration of EGF binding sites seems to be around the 14th day of gestation; however, this is only a rough estimate because different stages of maturation

are encountered within one uterus. Amnion tissue showed the greatest binding activity (expressed per mg protein), followed by, in decreasing order, lungs, trophoblast, limbs, visceral yolk sac (VYS), liver, brain, parietal endoderm and heart.

In the standard assay described by us previously⁸, tissues were exposed for 90 min at room temperature to tracer amounts of radioiodinated EGF in the presence or absence of a 500-fold excess of unlabelled EGF. Although nonspecific binding of whole tissue was higher than that found for cultured cells⁸ (see Fig. 1c and Table 1 legend), specific binding was still readily measurable. Three processes seem to contribute to these values: (1) specific binding (that which is displaced by a large excess of unlabelled EGF); (2) endocytosis of receptor–ligand complexes (the inclusion of 0.1% sodium azide in the binding medium reduced the bound counts by 21% in amnion and by 25% in VYS and brain); (3) nonspecific binding and endocytosis detected by ¹²⁵I-EGF counts which are not displaced in the presence of cold EGF.

When cells were isolated from 15-day embryonal tissues using trypsin and then grown in monolayer cultures, the nonspecific binding was <2% of the specifically bound counts. This suggests that for whole tissues some ¹²⁵I-EGF is trapped, whereas trypsin-treated cells have better access to the binding medium as well as a less 'sticky' surface.

An additional difficulty in using whole tissues is that of the accessibility of the tracer to the deeper-lying cells. We therefore made crude cell membrane preparations of each tissue and measured binding to aliquots of these preparations (Table 1). Although we used fewer samples than for whole tissue, the results with membrane preparations confirmed the whole tissue results.

In the experiments described above, parietal endoderm (cells plus basement membrane) showed low but significant binding of ¹²⁵I-EGF. However, this tissue is difficult to dissect entirely free of trophoblast cells, which bind higher levels (see Table 1). In addition, autoradiographic studies showed that Reichert's membrane can bind EGF. Therefore, we prepared single-cell suspensions of parietal endoderm cells by incubating yolk sacs with trypsin (0.25% w/v) and EDTA (0.1 mM) in phosphate-buffered saline (PBS) and allowed these cells to attach to plastic dishes (10⁵ cells per 35-mm dish). Cells were incubated in

Table 1 *In vitro* binding of ^{125}I -EGF to mouse embryonic tissues

Tissue	Day of gestation			
	11th and 12th	13th and 14th	15th and 16th	17th and 18th
Placenta membranes*	0.46 (3)	1.33 (37)	2.70 $\pm$ 2.1 (8)	0.35 (4)
Trophoblast	0.52 $\pm$ 0.05 (5)	2.34 $\pm$ 0.07 (7)	—	—
PYS endoderm*	0.25 $\pm$ 0.02 (4)	0.32 $\pm$ 0.02 (10)	—	—
VYS	1.82 $\pm$ 0.02 (5)	0.73 $\pm$ 0.03 (11)	1.02 $\pm$ 0.04 (12)	0.88 $\pm$ 0.03 (5)
Amnion	14.6 $\pm$ 5.9 (5)	21.5 $\pm$ 5.3 (11)	15.7 $\pm$ 3.5 (12)	9.1 $\pm$ 4.0 (5)
Amnion membranes*	ND	78.7 (2)	67.3 (1)	13.4 (1)
Limbs	—	1.7 $\pm$ 1.2 (4)	2.2 (1)	1.26 (1)
Lungs	—	4.26 (2)	0.44 $\pm$ 0.16 (4)	0.86 (1)
Lung membranes*	—	2.2 (1)	1.08 $\pm$ 0.41 (8)	0.66 $\pm$ 0.19 (5)
Liver	—	0.66 $\pm$ 0.18 (5)	0.08 (4)	0.23 (3)
Liver membranes*	—	0.96 $\pm$ 0.82 (4)	0.44 $\pm$ 0.24 (8)	0.52 (3)
Brain	—	0.48 $\pm$ 0.13 (4)	1.06 (3)	0.7 (1)
Brain membranes*	—	6.8 (3)	0.36 $\pm$ 0.10 (8)	0.43 $\pm$ 0.02 (5)
Heart	—	0.2 (3)	0.9 (2)	ND
Heart membranes*	—	ND	1.63 $\pm$ 0.97 (7)	2.23 (3)

EGF (Collaborative Research) was iodinated and purified by passage through Sephadex G-50 as described by Rees *et al.*⁸. Embryos were dissected in PBS (solution A)²² and washed in the same solution. Individual or pooled tissues were incubated in 24-well microtitre plates or in plastic tubes with 1 ml 'binding' buffer. Binding buffer contained 1 mg ml⁻¹ bovine serum albumin in Earle's balanced salt solution and 15 mM HEPES adjusted to pH 7.4 with 1 M NaOH. Iodinated EGF was added to 8×10^{-11} M (0.4 ng, 70,000 c.p.m.). After gentle agitation for 90 min at room temperature, tissues were washed three times in the same buffer at 0 °C (minus EGF) and twice in ice-cold PBS. Control samples were incubated similarly except for the addition of unlabelled EGF at 500 times the concentration of the labelled form. Each washed tissue was solubilized in 0.5–1.0 ml 0.2 M NaOH at 37 °C for 30 min and counted for 1 min in a Wallac γ -counter. Samples were assayed for protein by the Coomassie blue method²³. Similar assays for ^{125}I -EGF binding were carried out on crude membrane preparations. Various tissues were homogenized in glass/glass 1.0 ml homogenizers with 0.5 ml 0.01 M Tris-HCl, 0.01 M NaCl and 3 mM MgCl₂, pH 8.0. The homogenates were centrifuged at 100,000g for 30 min and the pellets resuspended in the binding buffer described above. Aliquots were tested for specific binding of ^{125}I -EGF and for protein content as above. The membranes were washed by suspension and centrifugation at 10,000g for 10 min or 5,000g for 20 min. Nonspecific binding of ^{125}I -EGF to embryonic tissues varied from 20% of total counts bound in amnion to 51.7% in placenta; in the crude membrane preparations, nonspecific binding varied from 26% for brain to 76% in limbs. For each tissue in each experiment the results were expressed as c.p.m. ^{125}I -EGF bound per mg protein minus c.p.m. ^{125}I -EGF bound in the presence of 500-fold excess of EGF per mg tissue protein divided by the total c.p.m. added to the incubation mixture and multiplied by 100. At least six such samples of each tissue were examined in each experiment (three reaction samples and three controls). ND, not done; PYS, parietal yolk sac; VYS, visceral yolk sac. The number of experiments is shown in parentheses.

*Crude plasma membrane preparations. parietal yolk sac endoderm cells plus Reichert's membranes are also contaminated with trophoblast cells.

Fig. 1 Autoradiographs of 14-day mouse embryo tissues to show ^{125}I -EGF binding and ^3H -TdR incorporation into the nucleus. For binding studies, tissues were incubated for 90 min: *a, d*, in 0.3 ml binding medium containing 250,000 c.p.m. ^{125}I -EGF; *b, e*, the same plus 0.1% sodium azide; *c*, in 0.3 ml binding medium containing 250,000 c.p.m. ^{125}I -EGF plus 500 ng unlabelled EGF. The tissues were washed six times in binding buffer and were fixed in phosphate-buffered 1% (w/v) formaldehyde for 4–6 h at room temperature. Tissues were washed, dehydrated and embedded in paraffin wax, sectioned at 8 μm and slices were dried on to slides. After dewaxing, slides were exposed to Kodak AR-10 stripping film for 6 weeks before being developed. *a–c*, VYS; *d, e*, amnion. Photographic reproduction of these sections is difficult and therefore diagrams have been drawn above the autoradiographs to indicate the relative approximate density of grains. Note that the VYS endoderm still takes up labelled EGF in the presence of a large excess of cold EGF (*c*). This nonspecific labelling probably occurs partly by endocytosis, a process which is characteristic of VYS endoderm²⁴. *d*, Shows that one cell layer of the amnion (tentatively identified as mesoderm) can bind greater amounts of EGF than the other cell layer. *f*, Autoradiograph of VYS incubated *in vitro* with EGF (5 ng ml⁻¹) for 16 h and then with ^3H -TdR (1 μCi ml⁻¹) for 4 h before being fixed and processed as above. End, endoderm; ect, ectoderm; mes, mesoderm. Scale bar, 20 μm .

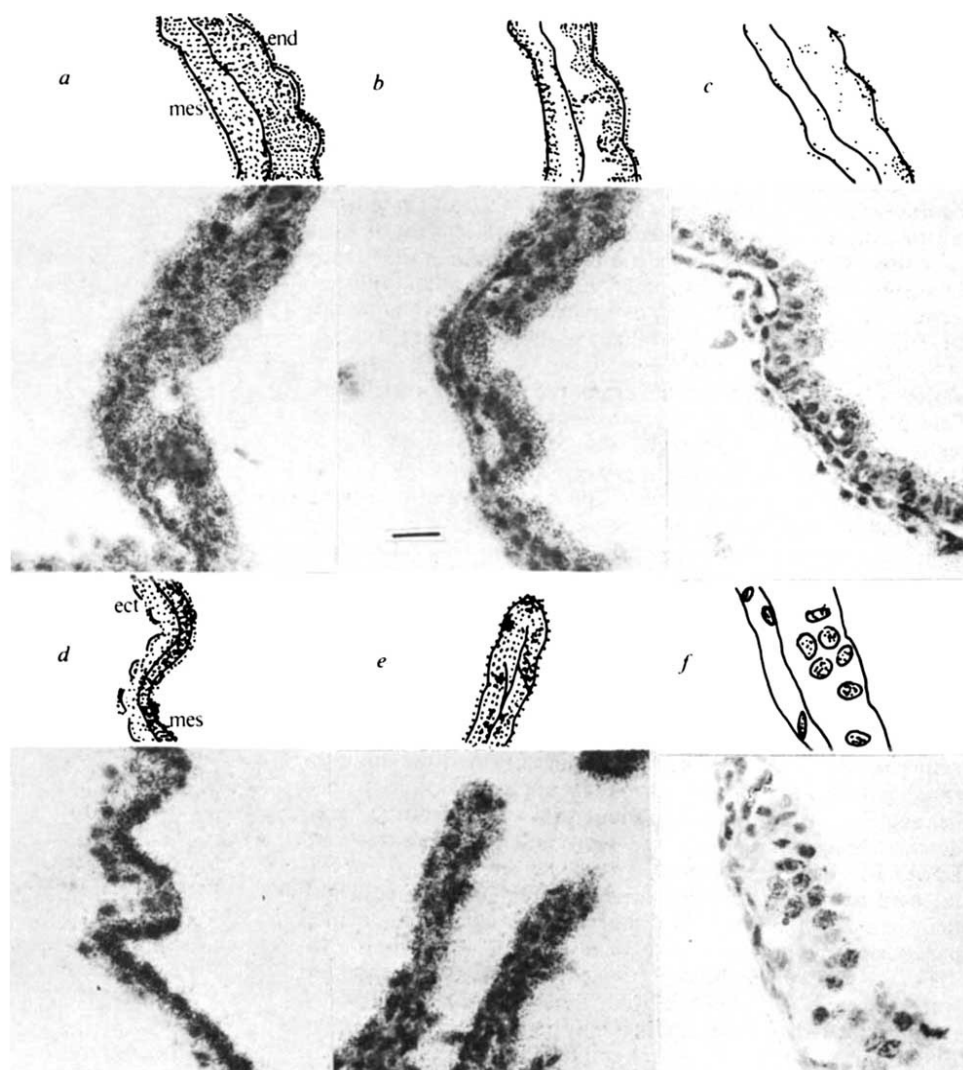


Table 2 Incorporation of ^3H -TdR into extra-embryonic tissues *in vivo*

	c.p.m. per mg protein ($\times 10^{-3}$)		
	Placenta	VYS	Amnion
EGF injected	1.6 (7)	15.0 (7)	133.0 (4)
PBS injected	5.6 (8)	443.0 (7)	304.8 (5)
Ratio PBS/EGF	3.5	29	2.3

The values represent averages of three different experiments on four different females, with the number of embryos examined in parentheses. Experimental method is described in the text and the tissues were processed as described in Fig. 2 legend.

medium which had been 'conditioned' by PYS-2 teratocarcinoma cells⁹ and filtered through 0.2- μm Millipore filters. Parietal endoderm cells attached well in these conditions and were incubated for 24 h to allow them to recover from any loss of trypsin-sensitive EGF-binding activity. Similar cell preparations were made from amnions to act as a standard cell type with high binding activity. Binding assays⁸ on cells attached to dishes were performed essentially as described for suspended tissues. Little or no ^{125}I -EGF binding was detected in parietal endoderm, whereas 4.1% specific binding per 10^6 cells was detected in amniotic cells.

Visceral yolk sac consists of two layers, endoderm cells underlain by basement membrane, and mesoderm consisting of several different kinds of cells including endothelial, fibroblastic and haematopoietic cells. Amnion has an ectodermal and a mesodermal layer. We examined the distribution of EGF-binding activity in the layers of VYS and amnion by autoradiographs prepared as described in Fig. 1 legend, and found that the greatest concentration of specific labelling occurred in the endoderm of the VYS (Fig. 1a-c) and in one of the two layers of the amnion (Fig. 1d), although all cell types had binding activity.

EGF has a wide range of effects on different cell types *in vitro*^{5,10-15}. The response of a target cell is usually determined experimentally through the mitogenic activity of EGF which may be measured by increased uptake of tritiated thymidine into DNA. Using assays described in Fig. 2 legend, we measured the incorporation of ^3H -thymidine (TdR) into DNA of explants of several different embryonic tissues, and found that for at least three tissues there was a dose-dependent response to EGF, with a maximal effect at $\sim 5 \text{ ng ml}^{-1}$ EGF. Autoradiographs of sections of tissues from 13-day embryos clearly showed that the ^3H -TdR incorporation was localized in the nuclei (Fig. 1f).

As reported previously¹⁶, we found that higher concentrations of EGF caused declining rates of incorporation (Fig. 2). This phenomenon has been attributed to receptor 'down-regulation' due to exposure of cells to high concentrations of EGF. It is believed to be a mechanism for reducing a cell's responsiveness to further stimulation by EGF and occurs by persistent internalization and degradation of EGF-receptor complexes. As we show below, such a process could also be occurring *in vivo*.

The uteri of anaesthetized female mice were exposed and 10 ng EGF per embryo were injected into the amniotic cavities. In another female, control embryos were injected with the same volume (2 μl) of PBS, and 16-18 h later, the uteri were again exposed and the embryos injected with 2 μl ^3H -TdR (1 μCi , specific activity 24 Ci mmol^{-1}) plus 10 ng EGF for experimental embryos. The mice were killed 90 min later and the embryos examined. Only embryos normal in appearance and with a beating heart were dissected (about 15% of the embryos were dead). Tissues were processed for counts incorporated into DNA. In eight tissues from 13-14-day embryos, the PBS-injected embryos had incorporated more ^3H -TdR into DNA than had the EGF-injected embryos, a result which seemed paradoxical in view of the *in vitro* results.

For three extra-embryonic tissues the counts incorporated were reduced by 2.3- to 29-fold those in the EGF-injected embryos (Table 2). These findings may be simply explained if the EGF injected was toxic, but general toxicity is unlikely because of the variable ratio of reduced incorporation in different tissues

and because even when massive doses of EGF were injected, similar results were obtained. Also arguing against toxicity are the parallel results achieved *in vivo* and *in vitro* at similar concentrations of EGF and ^3H -TdR. Instead, we hypothesize that EGF receptors may be down-regulated by prolonged exposure of the embryo to exogenous EGF.

EGF is present in mouse serum, urine and milk¹⁷ and also occurs in human amniotic fluid (at 16-18 weeks gestation, 0.9-8.8 ng ml^{-1} , ref. 18). EGF-like material is extractable from mouse fetuses at levels ranging from 0.6 to 15 ng per 14-day embryo⁵ although its level in amniotic fluid has not been reported. In our experiments the addition of $\geq 10 \text{ ng}$ of EGF per embryo seems to be sufficient to down-regulate the receptors. Such a conclusion is supported by the data in Fig. 2.

EGF has several roles other than mitogenic effects and we therefore cannot expect a simple relationship between EGF binding and thymidine incorporation, although for fibroblasts and many other cell types *in vitro*, this relationship largely holds. For embryonic tissues, it can also be speculated that EGF stimulates the differentiation of these tissues as it does in lung development in sheep¹⁹, rabbits⁷ and chicks²⁰.

In conclusion, most embryonic tissues bind EGF. Binding correlates with ^3H -TdR incorporation *in vitro* and evidence is provided that down-regulation of EGF receptors occurs *in vitro* and perhaps *in vivo*. These are some of the essential criteria which identify specific receptors for EGF²¹. The presence of EGF binding activity may serve to distinguish between extra-embryonic endoderm cell types, as visceral endoderm but not parietal endoderm binds EGF. This property will be useful for characterizing the endoderm cell products of teratocarcinoma EC cells.

We thank Professor R. L. Gardner for inviting J.B.W. to work in his laboratory, and Drs C. F. Graham and D. A. Mercola for useful discussions and comments on the script. This work was supported by grants to Dr C. F. Graham from the MRC, the Cancer Research Campaign and NIH grant 08293.

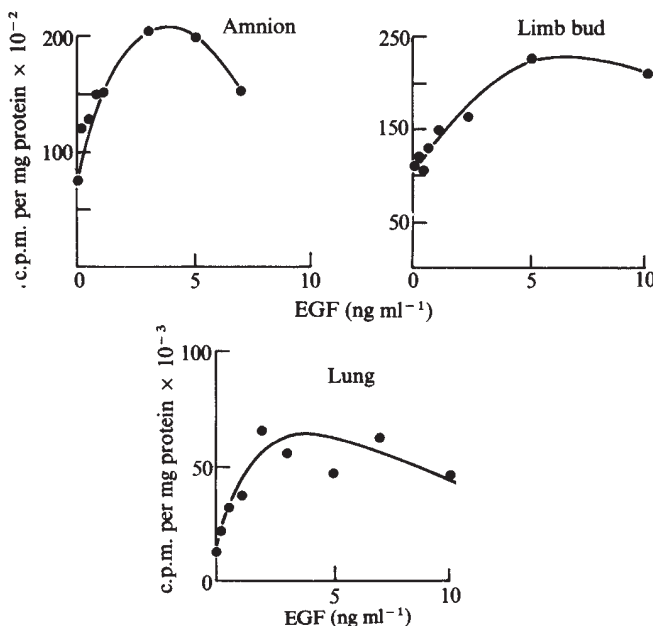


Fig. 2 The stimulation of DNA synthesis by EGF in mouse embryo tissues incubated *in vitro*. Tissues from 15-day embryos were dissected and incubated in culture media containing 10% dialysed fetal calf serum and concentrations of EGF varying from 0 to 10 ng ml^{-1} and 1 $\mu\text{Ci ml}^{-1}$ ^3H -TdR (24 Ci mmol^{-1}) for 16 h at 37 °C in a gassed incubator. The tissues were washed three times in PBS and then unincorporated label was removed by solubilization of tissues in 0.5 ml 0.5 M NaOH and subsequent precipitation by 0.5 ml 20% (w/v) trichloroacetic acid. After washing twice by resuspension and centrifugation the pellet was assayed for protein content and an aliquot tested for incorporated tritium radioactivity. The results are expressed as c.p.m. incorporated per mg protein in the sample for each EGF concentration. The graph represents one of three similar experiments for three different tissues but other tissues gave similar results.

Received 24 November 1980; accepted 14 April 1981.

1. Carpenter, G. & Cohens, S. A. *Rev. Biochem.* **48**, 193–216 (1979).
2. Hollenberg, M. D. *Vitam. Horm.* **37**, 69–110 (1979).
3. Adamson, E. D. & Rees, A. R. *Molec. cell. Biochem.* **34**, 129–152 (1981).
4. Cohen, S. *Dev. Biol.* **12**, 394–407 (1965).
5. Nexo, E., Hollenberg, M. D., Figueroa, A. & Pratt, R. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2782–2785 (1980).
6. O'Keefe, E., Hollenberg, M. D. & Cuatrecasas, P. *Archs Biochem. Biophys.* **164**, 518–526 (1974).
7. Catterton, Z. *et al. Pediat. Res.* **13**, 104–108 (1979).
8. Rees, A. R., Adamson, E. D. & Graham, C. F. *Nature* **281**, 309–311 (1979).
9. Lehman, J. M., Speer, W. C., Swartzendruber, D. E. & Pierce, G. B. *J. cell. Physiol.* **84**, 13–28 (1974).
10. Johnson, L. K., Baxter, J. D., Vlodavsky, I. & Gospodarowicz, D. *Proc. natn. Acad. Sci. U.S.A.* **77**, 394–398 (1980).
11. Tonelli, Q. J. & Sorof, S. *Nature* **285**, 250–252 (1980).
12. Covelli, I., Rossi, R., Mozzi, R. & Frati, L. *Eur. J. Biochem.* **27**, 225–230 (1972).
13. Carpenter, G., King, L. & Cohen, S. *Nature* **276**, 409–410 (1978).
14. Barnes, D. & Colowick, S. P. *J. cell. Physiol.* **89**, 633–640 (1976).
15. Hollenberg, M. D. & Cuatrecasas, P. *J. biol. Chem.* **250**, 3845–3853 (1975).
16. Fox, F. & Das, M. J. *supramolec. Struct.* **10**, 199–214 (1979).
17. Bynny, R. L., Orth, D. N., Cohen, S. & Doyno, R. S. *Endocrinology* **95**, 776–782 (1974).
18. Ladda, R. L., Bullock, L. P., Gianopoulos, T. & McCormick, L. *Analyt. Biochem.* **93**, 286–294 (1979).
19. Sundell, H. *et al. Pediat. Res.* **9**, 371A (1975).
20. Goldin, G. V. & Opperman, L. A. *Eur. J. cell. Biol.* **22**, 388A (1980).
21. Hollenberg, M. D. & Cuatrecasas, P. in *The Receptors* Vol. I (ed. O'Brien, R. D.) 193–213 (Plenum, New York, 1979).
22. Dulbecco, R. & Vogt, M. *J. exp. Med.* **99**, 167–182 (1954).
23. Sedmak, J. J. & Grossberg, S. E. *Analyt. Biochem.* **79**, 544–552 (1977).
24. Wild, A. E. *Phil. Trans. R. Soc. B271*, 395–410 (1975).

(–)Maackiain acetate specifically inhibits different forms of aryl hydrocarbon hydroxylase in rat and man

H. V. Gelboin*, Donna West*, E. Gozukara*, S. Natori†, Minako Nagao‡, & T. Sugimura‡

* Laboratory of Molecular Carcinogenesis, National Cancer Institute, Bethesda, Maryland 20205, USA

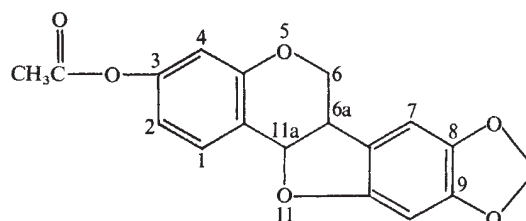
† National Institute of Hygiene Science, Kamiyoga, Setagayaku, Tokyo 158, Japan

‡ Division of Biochemistry, National Cancer Research Institute, Tsukiji, Chuo-ku, Tokyo, 104, Japan

The mixed-function oxidases¹, which contain different forms of cytochrome P450, are the critical enzymes which metabolize a wide variety of drugs, carcinogens and environmental chemicals, as well as steroids and other endogenous metabolites². The metabolism of these chemicals may result either in their detoxification or in their conversion to toxic, mutagenic and carcinogenic metabolites. Modulation of the mixed-function oxidase, aryl hydrocarbon hydroxylase (AHH), by either inhibitors or inducers strongly affects carcinogenesis induced by polycyclic hydrocarbons^{3,4}. We reported earlier that 7,8-benzoflavone is a potent inhibitor of the form of AHH that is induced by polycyclic hydrocarbons but either weakly inhibits or stimulates the form of AHH derived from the livers of phenobarbital-treated or untreated rats^{5,6}. We now report that the pterocarpan (–)maackiain acetate exhibits unique specificity that is the reverse of 7,8-benzoflavone, that is, it has strong inhibitory activity towards the form of AHH found in human liver or in untreated or phenobarbital-treated rats and weak inhibitory or stimulatory activity towards the AHH induced by polycyclic hydrocarbons in rat tissues or by cigarette smoking in human placenta. Thus, we conclude that (–)maackiain acetate and 7,8-benzoflavone inhibit different forms of AHH.

The structure of (–)maackiain acetate is shown below. (–)Maackiain is found in leguminous plants^{7,8} and was isolated from the heartwood constituents of *Sophora japonica* L. and converted to the acetate as described by Sugimoto^{7,9}. The physical characteristics of the material were as described previously⁷.

Three to five male Sprague–Dawley rats (120–160 g) were injected intraperitoneally with either 1.3 mg methyl-



(–) Maackiain acetate

cholanthrene per 0.3 ml corn oil per day for 3 successive days or 80 mg phenobarbital per 80 kg body weight per day for 3 days. The rats were killed and microsomal preparations prepared from liver, kidney and lung by published procedures⁶. Human liver was obtained within 6 h after the death of an accident victim and microsomes were prepared by a similar method to that described for rat liver. Placental tissue was obtained immediately post partum from women who had smoked at least 20 cigarettes per day for an extended period. The placentas were chilled on ice and microsome pellets obtained by essentially the same procedure as for rat liver. Protein content was determined by the method of Lowry *et al.*¹⁰. AHH activities of the various preparations were determined during a 10-min incubation with liver microsomes and for a 30-min incubation with microsomes from lung, kidney and placenta¹¹. The concentration of the substrate benzo(a)pyrene was 10^{-4} M in all experiments. Metabolite analysis of benzo(a)pyrene metabolism was analysed by HPLC using methods we have described previously¹².

Table 1 contrasts the effects of (–)maackiain acetate and 7,8-benzoflavone on AHH activity. At 10^{-4} M (–)maackiain acetate, a concentration equimolar to the substrate concentration, AHH activity in liver microsomes was inhibited by 95% in control rats and 76% in phenobarbital-treated rats, suggesting that the phenobarbital inducible enzymes may be rather less susceptible to (–)maackiain acetate inhibition than the enzyme forms in control rats. (–)Maackiain acetate either stimulated or had no inhibitory effect on the AHH of liver microsomes from methylcholanthrene-induced rats. In contrast, 7,8-benzoflavone strongly inhibits the AHH from methyl-

Table 1 Differential effects of (–)maackiain acetate and 7,8-benzoflavone on AHH of liver, lung, and kidney from normal, phenobarbital- and methylcholanthrene-treated rats

Expt	Tissue	In vivo treatment	Inhibitor (10^{-4} M)	AHH (pmol per min per mg protein)	% Control
1	Liver	Control	None	581	—
			MA	31	5
			7,8-BF	1,030	178
		Phenobarbital	None	862	—
			MA	210	24
			7,8-BF	1,300	151
		Methyl-cholanthrene	None	1,884	—
			MA	2,982	158
			7,8-BF	636	34
		Lung	None	8	—
2	Lung	Control	MA	3	37
			7,8-BF	7	87
		Methyl-cholanthrene	None	107	—
			MA	102	95
			7,8-BF	40	37
		Kidney	None	23	—
			MA	1	4
			7,8-BF	17	74
		Methyl-cholanthrene	None	101	—
			MA	56	55
			7,8-BF	39	39

Experiments 1 and 2 were performed on different sets of rats. MA, (–)maackiain acetate; 7,8-BF, 7,8-benzoflavone.

Table 2 Differential effects of (–)maackiain acetate and 7,8-benzoflavone on AHH of normal human liver and placenta from cigarette-smoking women

Tissue	Inhibitor (M)	AHH (pmol per min per mg protein)	% Control
Liver	None	11.0	—
	MA (10^{-4})	3.5	3.2
	MA (10^{-5})	5.5	50
	7,8-BF (10^{-5})	39.0	354
Placenta	None	15.6	—
	MA (10^{-4})	20.4	131
	MA (10^{-5})	20.6	132
	7,8-BF (10^{-4})	3.5	22
	7,8-BF (10^{-5})	8.1	39

cholanthrene-induced rats but has little or no inhibitory effect or stimulates the AHH from controlled or phenobarbital treated rats^{5,6,13,14}.

(–)Maackiain acetate and 7,8-benzoflavone also distinguish between two forms of the enzyme in lung and kidney. With lung microsomes, (–)maackiain acetate strongly inhibits AHH from control rats and does not inhibit the methylcholanthrene-induced enzyme, whereas 7,8-benzoflavone inhibits only the latter. The 7,8-benzoflavone inhibition of lung AHH from methylcholanthrene-induced rats has been reported to be only slightly greater than that from control rats¹³. With kidney microsomes, the (–)maackiain acetate inhibits both the control and methylcholanthrene-induced AHH but the latter to a much lesser extent. These data indicate that the composition of AHH present in the lung and kidney of untreated rats is different from that in the methylcholanthrene-inducible tissues.

Table 3 Characteristics of inhibition of (–)maackiain acetate and 7,8-benzoflavone in rat liver microsomes

	Inhibitor	Microsome source	
		Control	Pheno-barbital Methyl-cholanthrene
ID ₅₀	(–)MA	1×10^{-6} M	2×10^{-5} M
	7,8-BF	—	2×10^{-5}
K_m	None	7×10^{-6}	1.2×10^{-5}
	MA (10^{-5} M)	8×10^{-6}	—
	7,8-BF (10^{-5} M)	—	1.2×10^{-5}
V_{max}	None	1,720	1,000
	(–)MA (10^{-5} M)	200	—
	7,8-BF (10^{-4} M)	—	500
K_I	MA	(1.6×10^{-6})	—
	7,8-BF	—	8.8×10^{-6}

The ID₅₀ is the molar concentration that gives 50% inhibition and was calculated from direct measurement of inhibition at different concentrations of inhibitor. The K_I was calculated from Michaelis-Menten curves. V_{max} , units pmol per min per mg protein.

Table 2 shows the effects of (–)maackiain acetate and 7,8-benzoflavone on normal human liver, presumably uninduced for AHH, and the AHH of human placenta induced by heavy cigarette smoking, thus representing, respectively, control and polycyclic hydrocarbon-induced human tissues. (–)Maackiain acetate strongly inhibits human liver AHH and slightly stimulates the AHH of the induced placenta. AHH is essentially undetectable in the placentas of non-smoking women. In contrast, 7,8-benzoflavone caused a large stimulation of AHH in human liver and strongly inhibited inducible AHH of human placenta. A similar stimulation of control and phenobarbital-induced rat liver AHH^{13,14} and human liver AHH¹⁵ has been reported previously.

Table 3 shows some of the characteristics of (–)maackiain acetate and 7,8-benzoflavone with respect to inhibitory potency and effect on the kinetic characteristics of the AHH. The ID₅₀ of (–)maackiain acetate for control rat liver AHH is 10^{-6} M; that of 7,8-benzoflavone for methylcholanthrene-inducible rat liver AHH is 10^{-5} M. Each inhibitor reduces the V_{max} and has relatively little effect on the K_m of the AHH. Both inhibitors are non-competitive with substrate and (–)maackiain acetate is about 10-fold more potent for the control AHH than 7,8-benzoflavone for the polycyclic hydrocarbon-induced AHH.

Table 4 shows an HPLC analysis of the effect of (–)maackiain acetate on the formation of individual benzo(a)pyrene metabolites. Our earlier studies showed that 7,8-benzoflavone reduced the amount of each metabolite formed by methylcholanthrene-induced microsomes to about the same extent¹⁶. (–)Maackiain acetate drastically inhibits the formation of the 7,8-diol, 9,10-diol, 9-OH, 3-OH and 7-OH, but has a negligible effect on the formation of the 4,5 diol, suggesting that this metabolite may be formed by a different mixed-function oxidase from that metabolizing the formation of the other metabolites. The amount of quinones formed is reduced to a lesser extent than the major metabolites, which suggests either an alternative mixed-function oxidase for their formation or possibly their formation by a mechanism other than via the mixed-function oxidases.

The multiple forms of the cytochrome P450-containing mixed-function oxidases display different catalytic activities in the metabolism of various drugs^{1,17} and exhibit both substrate and product regio- and stereospecificity for the polycyclic hydrocarbon benzo(a)pyrene, a prototype carcinogen that is metabolized by this enzyme system¹². Metabolic pathway choice for a specific xenobiotic may thus be largely determined by the type and amount of the mixed-function oxidase engaging its metabolism.

Of a series of derivatives of 7,8-benzoflavone and other flavonoids, several were found either to inhibit the methylcholanthrene-induced AHH form only or to inhibit both the control and methylcholanthrene-induced forms^{13,18}. None of the many flavonoids tested exhibited the strong inhibitory activity specific for the control enzyme that is shown by (–)maackiain acetate. The unusual inhibitory specificity of (–)maackiain acetate may make it an extremely valuable tool for probing the enzymology and active catalytic sites of the different

Table 4 HPLC analyses of (–)maackiain acetate effect on benzo(a)pyrene metabolism in liver microsomes from control and MC-treated rats

Metabolite formed	Control microsomes			MC-microsomes		
	No inhibitor (pmol per min per mg protein)	(–)Maackiain acetate (pmol per min per mg protein)	% Control	No inhibitor (pmol per min per mg protein)	(–)Maackiain acetate (pmol per min per mg protein)	% Control
9,10-diol	33.2	6.0	18	572	608	106
4,5-diol	25.0	24.0	96	217	289	133
7,8-diol	8.6	0.7	8	326	368	113
9-OH	5.8	—	<1	241	386	160
7-OH	—	—	—	18	—	—
3-OH	222.1	17.8	8	1,261	1,471	116
Quinones	39.3	98.0	249	460	455	101
Total	334.0	146.5	43	3095	3577	729

forms of cytochrome P450, and the factors that determine metabolic pathway choice, and may thus help us to understand the relationship between drug and carcinogen activation and detoxification. It is also possible that individual or combinations of potent inhibitors of the mixed-function oxidases, such as (-)-maackiain acetate and 7,8-benzoflavone, may be powerful natural inhibitors or modulators of chemical carcinogenesis.

Received 15 December 1980; accepted 6 April 1981.

1. Coon, M. J. & Vatsis, K. P. in *Polycyclic Hydrocarbons and Cancer* Vol. 2. (eds Gelboin, H. V. & Ts'o, P. O. P.) 335–360 (Academic, New York, 1978).
2. Conney, A. H. *Pharmac. Rev.* **19**, 317–366 (1967).
3. Gelboin, H. V. *Physiol. Rev.* **60**, 1107–1166 (1980).
4. Wattenberg, L. W. *Cancer Res.* **26**, 1520–1526 (1966).
5. Diamond, L. & Gelboin, H. V. *Science* **166**, 1023–1025 (1969).
6. Wiebel, F. J., Leutz, J. C., Diamond, L. & Gelboin, H. V. *Archs Biochem. Biophys.* **144**, 78–86 (1971).
7. Sugimoto, H. *Bull. chem. Soc. Japan* **39**, 1529–1534 (1966).
8. Shibata, S. & Nishikawa, Y. *Chem. pharm. Bull., Tokyo*, **11**, 167–177 (1963).
9. Sugimoto, H. *J. org. Chem.* **24**, 1655–1662 (1959).
10. Lowry, O. H., Rosebrough, N. A., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265–275 (1951).
11. Nebert, D. W. & Gelboin, H. V. *J. biol. Chem.* **243**, 6242–6261 (1968).
12. Deutsch, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 3123–3127 (1978).
13. Wiebel, F. J., Gelboin, H. V., Buu-Hoi, N. P., Stout, M. G. & Burnham, W. S. in *Chemical Carcinogenesis Pt A* (eds Ts'o, P. & Di Paolo, J. A.) 249–270 (Dekker, New York, 1974).
14. Wiebel, F. J. & Gelboin, H. V. *Biochem. Pharmac.* **24**, 1511–1515 (1975).
15. Kapitlnik, J., Poppers, P. J., Buening, M. K., Fortner, J. G. & Conney, A. H. *Clin. Pharmac. Ther.* **22**, 475–485 (1977).
16. Selkirk, J. K., Croy, R. G., Roller, P. P. & Gelboin, H. V. *Cancer Res.* **34**, 3474–3480 (1974).
17. Ryan, D. E., Thomas, P. E., Korzeniowski, D. & Levin, W. J. *J. biol. Chem.* **254**, 1365–1374 (1979).
18. Nesow, S. J. *med. Chem.* **22**, 1244–1247 (1979).

Idiosyncratic platelet responses to human tumour cells

Eva Bastida, Antonio Ordinas & G. A. Jamieson

American Red Cross, Blood Services Laboratories, Bethesda, Maryland 20014, USA

Platelets are thought to be involved in two important steps in the dissemination of cancer: (1) in their interaction with cells released from the primary tumour to form a mixed platelet–tumour cell embolus and (2) in the subsequent attachment or entrapment of this embolus at the vessel wall and the resulting extravasation of the tumour cell and growth of the secondary tumour¹. We have recently used the Baumgartner perfusion chamber² to evaluate the attachment phase of platelet–tumour cell interactions at subendothelial surfaces³. We have found that tumour cells which caused platelet aggregation (Hut 20 line, large cell carcinoma of the lung) were incorporated into mixed platelet–tumour cell thrombi on basement membrane or collagen microfibrils, while those tumour cells that did not cause platelet aggregation (A549 line, epithelial lung carcinoma), were not incorporated in this way. We have now found differences between different donors in the ability of their platelets to aggregate in response to various human tumour cell lines. These different individual susceptibilities seem to be largely independent of the cell line involved and may have implications in the evaluation of mechanisms of human neoplastic disease.

The primary interaction of platelets and tumour cells has usually been studied by aggregometry^{4–7} and a rough correlation has been established between the ability of tumour cells to aggregate platelets and their ability to induce lung metastases in mice⁴. The basis for this aggregation by tumour cells has not been determined but has generally been ascribed as being due to the release of ADP from platelets^{4,5}, to the thromboplastic activity of tumour cells⁸ or to a lipoglycoprotein complex of the tumour cell surface⁷. This complex may be shed as microvesicles⁴ and may be identical to the platelet aggregating material (PAM) which has been isolated from mouse tumour cells⁶. However, PAM cannot be isolated from tumour cells in all cases⁷ and other mechanisms may be involved for other cell types.

Chelation of Ca^{2+} by collection of blood into citrate anti-coagulant results in the irreversible loss of the ability of platelets to interact with tumour cells. In our experiments, platelet-rich plasma was prepared from blood collected in sodium heparin (5 U ml^{-1}) from healthy volunteers who had taken no drugs for the previous 7 days: platelet counts ($\sim 300,000$ per μl) and recoveries ($\sim 80\%$) were normal. Aliquots of washed tumour cells from each cell line were then added, to final concentrations given in Table 1, to an aliquot of platelet-rich plasma from each donor in the aggregometer (Chronolog, Broomall, Pennsylvania) and the extent of aggregation determined. Aggregation induced by $10 \mu\text{M}$ ADP was taken as 100% response. After culture of each cell line in standard conditions, cells were collected by decanting the culture medium, and the monolayers were washed twice with Hank's balanced salt solution (HBSS) followed by treatment for 5 min with HBSS free of Ca^{2+} and Mg^{2+} and containing 4 mM EGTA. The cell suspension was centrifuged at $800g$ for 10 min, the supernatant solution removed for testing and the cell pellets were washed twice with a solution of HBSS free of Ca^{2+} and Mg^{2+} but containing bovine serum albumin and 0.25 mg ml^{-1} apyrase. Finally they were resuspended in the same solution but without apyrase. Cells were counted in a haemocytometer and viability determined by exclusion of Trypan blue. The range of viable cells was between 90 and 97%. None of the cell lines showed contamination with fibroblasts and the effects observed were not affected by treatment of the cultures with collagenase.

The results are summarized in Table 1. Individual donors were examined 2–4 times over a period of 4 months. Results were constant with regard to aggregation response in comparison with the $10 \mu\text{M}$ ADP control and lag times did not vary by more than $\pm 10 \text{ s}$ in repeat experiments. At a concentration of 10^5 cells ml^{-1} the U87MG (glioblastoma) line caused the aggregation of platelet-rich plasma (PRP) from each of the 16 donors examined. With Hut 20 (large cell lung carcinoma) cells at this concentration, six of the donors showed full platelet aggregation response with normal lag time (2–3 min), while the lag time to the onset of aggregation was prolonged in two other donors (5 and 9) and six failed to show aggregation. At 10^6 cells ml^{-1} , the Hut 28 (mesothelioma) line gave no aggregation with the same six donors who had shown zero aggregation response with Hut 20 at 10^5 ml^{-1} , and also with the two donors (5 and 9) who had shown a prolonged lag time to onset of aggregation. Donor 10, who had shown full aggregation with normal lag time with Hut 20 cells, did not give an aggregation response with HT 29 (adenocarcinoma). All donors showed a full aggregation response with Hut 20 at 10^6 cells ml^{-1} . For Hut 28 and SKNMC (neuroblastoma), the aggregation threshold was not reached until 5×10^6 cells ml^{-1} but, at this concentration, the same nine donors did not show an aggregation response, as was the case with the HT 29 line at 10^6 cells ml^{-1} . None of the donors responded to A549 (epithelial lung carcinoma) at tumour cell concentrations as high as 10^7 ml^{-1} .

The maximum aggregation response was not the same for each cell line: the U87MG, Hut 20 and HT 29 lines gave maximum responses, that is, equal to that observed with $10 \mu\text{M}$ ADP, whereas the responses of the SKNMC and Hut 28 lines were 60 and 50% respectively, of the ADP value. However, these values were reproducible and probably reflect differences in the sizes of the platelet–tumour cell aggregates in these two cases.

Donor 7 was unusual in showing reduced aggregation ($\sim 40\%$) with U87MG at 10^5 cells ml^{-1} with a prolonged lag time of 4.2 min and there was no response with any other cell line.

These results show, not surprisingly, that certain tumour cell lines do not cause platelet aggregation (A549) whereas others (U87MG) cause aggregation at all tumour cell concentrations examined. However, for several lines, there are critical tumour cell concentrations within which donors can be grouped into responders and non-responders. More significantly, this division seems to be largely independent of the particular tumour cell line examined. The tumour cell lines used here show different

Table 1 Aggregation responses of individual donors to six human tumour cell lines

Donor	Cell line U87MG (10^5 ml^{-1})		Hut 20 (10^5 ml^{-1})		HT 29 (10^6 ml^{-1})		Hut 28 ($5 \times 10^6 \text{ ml}^{-1}$)		SKNMC ($5 \times 10^6 \text{ ml}^{-1}$)		A549 (10^7 ml^{-1})	
	%	T	%	T	%	T	%	T	%	T	%	T
1	100	2.7	100	2.8	100	2.2	50	4.0	60	3.0	0	—
2	100	2.2	100	2.6	100	3.0	50	4.5	60	3.2	0	—
3	100	3.1	0	—	0	—	0	—	0	—	0	—
4	100	3.5	0	—	0	—	0	—	0	—	0	—
5	100	2.6	90	6.9	0	—	0	—	0	—	0	—
6	100	2.0	100	2.0	100	4.1	50	3.5	60	2.7	0	—
7	40	4.2	0	—	0	—	0	—	0	—	0	—
8	100	2.2	100	3.0	100	2.5	50	3.2	60	3.0	0	—
9	100	3.0	100	4.4	0	—	0	—	0	—	0	—
10	100	3.2	100	3.0	0	—	0	—	0	—	0	—
11	100	2.8	100	3.0	100	3.0	50	3.0	60	2.5	0	—
12	100	2.3	100	2.9	100	3.0	50	3.8	60	3.0	0	—
13	100	2.9	100	3.0	100	3.5	50	4.2	60	3.5	0	—
14	100	3.5	0	—	0	—	0	—	0	—	0	—
15	100	2.6	0	—	0	—	0	—	0	—	0	—
16	100	3.5	0	—	0	—	0	—	0	—	0	—

% Aggregation responses compared with $10 \mu\text{M}$ ADP. T, lag time (min) from addition of tumour cells to onset of aggregation.

aggregation profiles and different effects are observed after treatment with apyrase, hirudin and phospholipases⁹, which suggests that several different mechanisms may be involved. However, the responses of individual donors to the various cell types observed here are remarkably constant. Of the four cell types in which different responses were observed (Hut 20, HT 29, Hut 28 and SKNMC), six of the donors (3, 4, 7, 14, 15 and 16) showed no response with any of the lines while two other donors (5 and 9) showed a response only with Hut 20 but with a prolonged lag time to the onset of aggregation. Only in the case of one donor (10) did there seem to be normal aggregation with Hut 20 and no aggregation with the other three lines. Further studies are necessary to determine whether a graduated response can be shown at low concentrations of U87MG cells although aggregation by this line seems to involve a unique mechanism.

These observations suggest a greater sensitivity of certain individual donors to platelet aggregation by human tumour cells in general, which may explain why certain types of tumours can be metastatic in one individual but not in another¹. The results may have more general significance in their application to the epidemiology and genetics of human neoplastic disease.

This work was supported in part by USPHS grants HL 14697, HL 20971 and Biomedical Research Support grant RR 05737. We thank Dr Jorgen Fogh for providing us with the U87MG, HT29 and SKNMC lines, and Dr Addi Gazdar for the Hut 20, Hut 28 and A549 lines.

Note added in proof: Dr Gazdar has since informed us that the Hut 20 line has now been identified as being of murine origin, based on cytogenetic analysis and isozyme phenotyping, while a human origin has been confirmed for the other tumour cell lines examined here. This removes discrepancies associated with the greater reactivity of Hut 20 cells with platelets and explains why 9 of the 16 donors (3–5, 7, 9, 10 and 14–16) failed to respond at the critical tumour cell concentration with all of the differentially reactive human lines HT 29, Hut 28 and SKNMC.

Received 19 November 1980; accepted 21 April 1981.

1. Sugarbaker, E. V. & Ketcham, A. S. *Semin. Oncol.* **4**, 19–32 (1977).
2. Baumgartner, H. R. *Microvasc. Res.* **5**, 167–179 (1973).
3. Marcum, J. M., McGill, M., Bastida, E., Ordinas, A. S. & Jamieson, G. A. *J. Lab. clin. Med.* **96**, 1046–1053 (1980).
4. Gasic, G. J., Kock, P. A. G., Hsu, B., Gasic, T. B. & Niewiarowski, S. *Z. Krebsforsch.* **86**, 263–277 (1976).
5. Holme, R., Oftebro, R. & Hovig, T. *Thromb. Haemostasis* **40**, 89–102 (1978).
6. Pearlstein, E., Cooper, L. B. & Karparkin, S. *J. Lab. clin. Med.* **93**, 332–344 (1979).
7. Hara, Y., Steiner, S. & Baldini, M. G. *Cancer Res.* **40**, 1212–1216 (1980).
8. Gastpar, H. *Thromb. Diath. haemorrh.* **42**, 291–303 (1970).
9. Bastida, E., Ordinas, A. & Jamieson, G. A. *Cancer Res.* (submitted).

Routes of conjugation in normal and cancerous tissue from human lung

Gerald M. Cohen*, Elizabeth M. Gibby & Rekha Mehta

Toxicology Unit, Department of Pharmacology, School of Pharmacy, University of London, 29/39 Brunswick Square, London WC1N 1AX, UK
Department of Biochemistry, University of Surrey, Guildford GU2 5XH, Surrey, UK

The selective toxicity of drugs leading to major advances in antibacterial chemotherapy has often resulted from the identification and exploitation of major biochemical differences between bacterial and mammalian species¹. Similar progress has not been made in cancer chemotherapy, partly due to a lack of suitable biochemical differences between normal and cancerous tissue other than in DNA synthesis, but also because of many other problems such as those of metastases and resistance, and the presence in tumours of cells at different states of the cell cycle. Here we report a major biochemical difference in the routes of conjugation between normal lung and tumour tissue from patients with lung cancer. Conjugation with glucuronic acid and sulphate constitute two of the most important pathways of metabolism of drugs, other foreign compounds and hormonal steroids^{2,3}. Using 1-naphthol as a model phenolic substrate, normal peripheral lung tissue formed almost exclusively the sulphate ester conjugate, 1-naphthyl sulphate, whereas tumour tissue from squamous carcinomas from the same patients formed predominantly the glucuronic acid conjugate, 1-naphthyl glucuronide. Such major biochemical differences may be exploitable in the design of selectively toxic cancer chemotherapeutic agents.

Peripheral lung tissue and tumours from patients with lung cancer were cultured within 3 h of surgical removal. When short-term organ cultures of peripheral human lung were incubated with [$1\text{-}^{14}\text{C}$]1-naphthol, the major metabolite formed was the sulphate conjugate, 1-naphthyl sulphate, with little or no glucuronic acid conjugate being formed (Table 1), which agrees with our previous results⁴. However, in the case of tumour tissue, the metabolism of 1-naphthol varied with the histological classification of the tumour (Table 1). Particularly striking was the finding that tumour tissue from all patients with

* To whom correspondence should be addressed.

Table 1 Conjugation of [1-¹⁴C]1-naphthol by short-term organ cultures of normal human lung and tumour tissue

Patient	Histological diagnosis	% Of total radioactivity recovered in medium as conjugates in the:			
		Lung		Tumour	
		Sulphate	Glucuronide	Sulphate	Glucuronide
VII	Poorly differentiated squamous cell	50.2	ND	0.7	39.8
XI	Poorly differentiated squamous cell	23.7	ND	1.0	8.3
X	Squamous cell	18.9	ND	ND	2.9
XXI	Very poorly differentiated squamous cell	11.5	1.0	1.1	6.9
XVI	Moderately well differentiated squamous cell	16.6	ND	0.1	7.7
XII	Undifferentiated large cell, possibly squamous	35.9	0.3	0.7	12.3
XVIII	Undifferentiated large cell, possibly squamous	11.3	2.0	0.7	13.4
XVII	Undifferentiated large cell	11.1	ND	3.4	6.2
XX	Undifferentiated large cell, possibly adenocarcinoma	13.7	4.4	2.0	3.9
IX	Poorly differentiated adenocarcinoma of large cell, undifferentiated	7.1	ND	ND	ND
XIV	Poorly differentiated adenocarcinoma (possibly metastatic from breast)	27.0	3.1	1.8	4.5
XXII	Small (large) cell anaplastic carcinoma	84.4	1.0	28.5	3.0
XXIII	Large (small) cell anaplastic carcinoma	61.0	3.7	22.0	8.1

Short-term organ cultures of macroscopically normal human lung or tumour tissue were cultured for 18 h at 37 °C either in Leibovitz L-15 medium as previously described⁴ or on gelatin sponge in 2.5 ml of CMRL-1066 medium (Biocult, Scotland) as described by Stoner *et al.*¹⁷. After 18 h the culture medium was changed for one containing [1-¹⁴C]1-naphthol (10–20 µM) (specific activity 19.2 Ci mmol⁻¹; Radiochemical Centre) and generally incubated for a further 90 min, but sometimes for another 24 h (patients XXII and XXIII). The amount of radioactivity in the medium at the end of the culture was 70–90% of the total. Aliquots of the media were analysed for 1-naphthol conjugates by TLC as described elsewhere⁴. The results are generally expressed as the mean values from two different explants. Considerable variability was occasionally observed, particularly for the tumour tissue probably due to its heterogeneity. However, the results always showed a similar pattern. Similar metabolite profiles were obtained when lung and tumour were cultured with naphthol: 100–1,000 µM or 5–100 µM respectively. ND, not detectable.

squamous cell carcinomas metabolized 1-naphthol predominantly to its glucuronic acid conjugate, with little or no sulphate conjugate being formed (Table 1). Limited results from other types of lung tumours prevent detailed analysis but show a greater variability in the routes of conjugation (Table 1). These results are supported by our recent studies with human lung cancer cell lines, when cultures derived from squamous cell carcinomas showed extensive formation of 1-naphthyl glucuronide with little sulphate conjugation⁵.

To confirm and extend these major differences observed for short-term organ cultures, the rates of conjugation of 1-naphthol by 10,000g supernatant fractions of normal lung or tumour tissue were also determined. In patients with sufficient tissue for study, rates of glucuronic acid conjugation were much greater than sulphate ester conjugation in the tumour tissue, whereas in normal lung the reverse was true (Table 2), thus confirming the differences observed using cultured explants.

It is important to stress that the present studies have used lung tissue but most human lung cancers arise from the bronchial epithelium⁶. This should not be a disadvantage as sulphate conjugation is quantitatively far more important than glucuronic acid conjugation for cultured human bronchus⁷. Due to the multiplicity of both UDP-glucuronosyltransferase² and sulphotransferase³ activities, it is difficult to predict the metabolic fate of different substrates from results obtained for one substrate. Nevertheless, with 1-naphthol as substrate, tumorigenesis in human lung is accompanied by the appearance of a conjugation pathway which has little or no activity in the host lung. These changes in conjugation may be affected by alterations in β -glucuronidase and aryl sulphatase, which may be elevated in tumours⁸. However, higher activity of β -glucuronidase in tumours would only minimize the changes noted here. Other tumours may also show similar changes in conjugation; a marked shift from sulphation to glucuronidation was observed in recent studies on cultures of human colon and colonic tumour tissue⁹. The relationship, if any, of these changes in conjugation to tumour development and progression remains obscure, although they may be connected to alterations in cell-surface glycosaminoglycans.

Various rodent hepatomas have shown similar increases in UDP-glucuronosyltransferase activity, accompanied by greatly decreased phenolsulphotransferase activity in a Reuber H-35 hepatoma^{10–12}. These results and those reported here showing increased activities of UDP-glucuronosyltransferase in both

human and rodent tumours, are in contrast to the generally reported absence or marked decreases in mixed function oxidase activities in both pre-neoplastic nodules and tumours^{13,14}. Such alterations in metabolism may have important implications in the successful chemotherapy of neoplastic disease. As some anticancer agents, for example, cyclophosphamide, require metabolism to exert their activity, the balance of oxidative and conjugating enzymes in the tumour will be one of the critical factors determining the amounts of reactive metabolites generated and thus the efficacy of the treatment.

Although conjugation of drugs is generally considered to be a detoxification process, there are examples of conjugation, including glucuronidation^{15,16}, which lead to more reactive metabolites. Based on the biochemical differences observed

Table 2 Rates of conjugation of [1-¹⁴C]1-naphthol by 10,000g supernatant fractions of normal human lung and tumour tissue

Patient	Rate of conjugation (nmol per min per g protein) in:			
	Lung		Tumour	
	With sulphate	With glucuronic acid	With sulphate	With glucuronic acid
VII	15.9	0.8	1.9	4.9
XII	11.6	3.4	2.7	16.5
XVII	85.2	3.6	7.5	27.9
XVIII	8.6	3.5	ND	29.1
XXI	6.9	1.7	0.1	100.2

Homogenates (50%) of normal human lung and tumour tissue were centrifuged at 10,000g for 30 min and the supernatant fractions either frozen overnight at –20 °C before analysis (for patients VII, XII and XXI) or used as soon as prepared (for patients XVII and XXII). UDP-glucuronosyltransferase activity was measured at 37 °C in 0.5 ml volume containing 2–6 mg protein, 1.5 mM UDP-glucuronic acid (Sigma), 4 mM MgCl₂, 20 µM [1-¹⁴C]1-naphthol (specific activity 19.2 Ci mol⁻¹) and 0.22 M Tris-HCl buffer, pH 7.4. Phenolsulphotransferase activity was measured at 37 °C in 0.5 ml volume containing 3–6 mg protein, 2 mM Na₂SO₄, 5 mM MgCl₂, 5 mM ATP (disodium salt, Sigma), 20 µM [1-¹⁴C]1-naphthol and 0.14 M Tris-HCl buffer, pH 7.4. Control incubations contained no enzyme fractions. Incubations were carried out for 10–60 min to ensure linearity of the reaction. Reactions were stopped by removal of 100-µl aliquots of the incubation media into tubes with an equal volume of methanol containing authentic 1-naphthol and metabolites, which were separated and analysed as described previously. ND, not detectable.

here, it may be possible to design drugs, activated by glucuronic acid conjugation but not by sulphate conjugation, which would be selectively toxic to the tumour cells.

We thank Mr Meredith-Brown and colleagues for the supply of tissues and Professors P. Grasso and B. Corrin for histological diagnosis of the tumours. This work was supported by a grant from the Cancer Research Campaign of Great Britain.

Received 19 January; accepted 1 April 1981.

- Albert, W. *Selective Toxicity: The Physico-chemical Basis of Therapy* (Chapman & Hall, London, 1973).
- Dutton, G. J., Wishart, G. J., Leaky, J. E. A. & Goheer, M. A. in *Drug Metabolism—From Microbe to Man* (eds Parke, D. V. & Smith, R. L.) 71–90 (Taylor & Francis, London, 1977).
- Dodgson, K. S. in *Drug Metabolism—From Microbe to Man* (eds Parke, D. V. & Smith, R. L.) 91–104 (Taylor & Francis, London, 1977).
- Mehta, R. & Cohen, G. M. *Biochem. Pharmacol.* **28**, 2479–2484 (1979).
- Gibby, E. M., Mehta, R., Ellison, M. & Cohen, G. M. *Biochem. Soc. Trans.* **9**, 111 (1981).
- Wynder, E. L. & Hecht, S. (ed.) *Lung Cancer, UICC tech. Rep. Ser.* **25**, 3–33 (International Union Against Cancer, Geneva, 1976).
- Autrup, H. et al. *Int. J. Cancer* **25**, 293–300 (1980).
- Fishman, W. H. & Anlyan, A. *J. Cancer Res.* **7**, 808–817 (1947).
- Cohen, G. M., Gibby, E. M., Grafström, R. & Harris, C. C. (in preparation).
- Lueders, K. K., Dyer, H. M., Thompson, E. B. & Kuff, E. L. *Cancer Res.* **30**, 274–279 (1970).
- Gessner, T. *Biochem. Pharmacol.* **23**, 1809–1816 (1974).
- Zimmerman, J. J., Gorski, J. P. & Kasper, C. B. *Drug Metab. Dispos.* **5**, 572–578 (1977).
- Strobel, H. W., Dignam, T. D., Saine, S. E., Fang, W. & Fennell, P. M. *Molec. cell. Biochem.* **22**, 79–91 (1978).
- Cameron, R., Sweeney, G. D., Jones, K., Lee, G. & Farber, E. *Cancer Res.* **36**, 3888–3893 (1976).
- Miller, E. C. et al. *Molec. Pharmacol.* **4**, 147–154 (1968).
- Mulder, G. J., Hinson, J. A. & Gillette, J. R. *Biochem. Pharmacol.* **27**, 1641–1649 (1978).
- Stoner, G. D. et al. *Lab. Invest.* **38**, 685–692 (1978).

Sequential appearance of Epstein-Barr virus nuclear and lymphocyte-detected membrane antigens in B cell transformation

D. J. Moss, A. B. Rickinson, L. E. Wallace & M. A. Epstein

Department of Pathology, University of Bristol Medical School, University Walk, Bristol BS8 1TD, UK

One of the central questions of Epstein-Barr (EB) virus biology concerns the nature of the asymptomatic carrier state^{1,2} whereby the virus, which *in vitro* transforms B cells into permanent lymphoblastoid cell lines with such efficiency^{3,4}, is harboured for life in the B lymphoid tissues of all previously-infected (seropositive) individuals⁵. It seems highly probable that this persistent infection is primarily controlled by memory T cells (cytotoxic precursors) specifically primed to the EB virus-induced lymphocyte-detected membrane antigen LYDMA⁶⁻⁹. Thus the degree to which the infection can progress without interruption in individual B cells of the immune host will depend upon the timing of LYDMA expression relative to that of other viral gene functions. We now demonstrate that during EB virus-induced B cell transformation *in vitro* LYDMA is expressed on the cell surface soon after the appearance of the virus-associated nuclear antigen EBNA¹⁰ and coincident with, but not dependent upon, the initiation of cellular DNA synthesis. This suggests that in the lymphoid tissues of the infected host, immunological control over virus-B cell interactions can be exercised early in the cycle at the very onset of virus-induced cell proliferation.

EB virus-specific memory T cells are present exclusively in seropositive individuals and can be reactivated *in vitro* by co-cultivation either with freshly-infected autologous B cells⁶ or with low numbers of cells from the autologous EB virus-transformed B cell line¹¹. Such reactivated T-cell preparations kill virus-infected cells through HLA-A and B antigen-restricted recognition of the target antigen LYDMA¹¹⁻¹⁵, a virus-induced

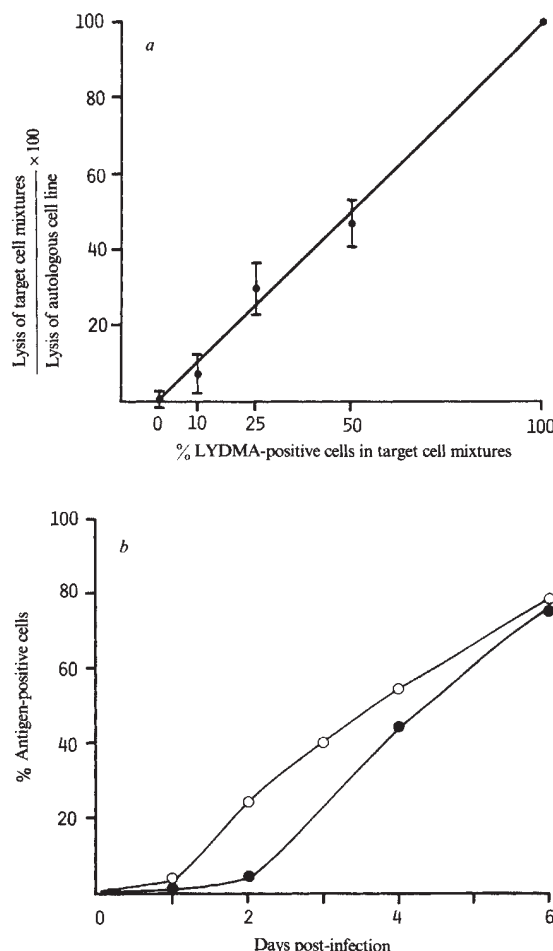


Fig. 1 *a*, Calibration curve of activity of effector T cells from donor D.M. Freshly prepared blood mononuclear cells from donor D.M. were seeded at 2×10^6 cells per 2-ml well (Linbro) with 2.5×10^4 cells of the autologous EB virus-transformed cell line. At 10 days later the E-rosette-forming subpopulation was prepared by centrifugation over Ficoll-Hypaque followed by lysis of the erythrocytes in NH_4Cl . Replicate ampoules of these effector T cells were stored in liquid nitrogen. Full details of these procedures are given elsewhere^{9,11}. Effector T cells from one such ampoule were tested in a preliminary chromium release assay for cytotoxicity against the autologous EB virus-transformed cell line (DM-EB), against EB virus-transformed cell lines from three HLA-A and -B antigen-unrelated donors M.O., J.W. and A.S. (MO-EB, JW-EB, AS-EB) and against two EB virus genome-negative cell lines K562 and HSB-2. Cytotoxicity was expressed in terms of the per cent specific lysis of the ^{51}Cr -labelled target cells observed following a 5-h incubation with effector T cells in V-shaped wells at an effector:target ratio of 20:1, as described elsewhere¹³. Levels of cytotoxicity were as follows: DM-EB, 52%; MO-EB, 8%; JW-EB, 2%; AS-EB, 3%; K562, 2%; HSB-2, 6%. Effector T cells from other ampoules of this preparation were then tested for cytotoxicity against autologous T-depleted (TD) mononuclear cells cultured without exposure to EB virus, against the autologous EB virus-transformed cell line and against artificial mixtures containing these two different types of target cells in varying proportions. The calibration curve shown was constructed by combining the results of three separate tests of this kind, carried out on days 1, 2 and 4 of an experiment with cells from donor D.M. For each artificial target cell mixture, the extent of lysis was on each occasion expressed as a percentage of that simultaneously recorded against the autologous target cell line itself and the mean percentage ($\pm$ s.d) from the three tests was then plotted against the proportion of LYDMA-positive cells (that is, cells of the autologous line) in that mixture. The absolute level of autologous cell line lysis recorded on these occasions was 55–70% specific isotope release; the absolute level of lysis of cultured autologous TD cells was always below 5%. *b*, Induction of EBNA-positive (○) and LYDMA-positive (●) cells in cultures of EB virus-infected TD cells from donor D.M. The percentage of EBNA-positive cells in each test population was obtained from counts on 1,000 cells. All appropriate controls were included within each EBNA-staining procedure, as described elsewhere²⁹. Only cells showing bright unequivocal staining with positive serum were scored as EBNA positive; preparations stained with negative serum showed virtually no cells with nuclear fluorescence. The percentage of LYDMA-positive cells in each test population was deduced from the extent of specific lysis in the chromium release assay (relative to the specific lysis of the autologous cell line itself) using the calibration curve of effector T-cell activity shown in *a*.

cell surface change whose existence was first invoked to explain the cellular response to primary EB virus infection seen in infectious mononucleosis patients¹⁶⁻¹⁸. Thus strong killing of the autologous, and of allogeneic HLA-A and B antigen-related, EB virus-transformed cell lines by *in vitro* reactivated T cells occurred in the absence of any significant lysis of (1) the corresponding cell lines from HLA-unrelated donors, (2) autologous pokeweed mitogen (PWM)-stimulated B lymphoblasts, or (3) EB virus genome-negative human haemopoietic cell lines such as K562 and HSB2 deliberately included as sensitive indicators of non-specific or natural killer-like activities which under other circumstances¹⁹⁻²² can arise in cultured T-cell populations (see Fig. 1 legend, Fig. 2 and refs 11, 13, 15).

Blood mononuclear cells from seropositive donors were co-cultivated with cells of the autologous EB virus-transformed cell line at a ratio of 80:1 and the E-rosette-forming subpopulation collected from these co-cultures 8 to 12 days later provided EB virus-specific cytotoxic T cell preparations which were stored in replicate aliquots in liquid nitrogen until required²³. In the present experiments these effector T cells have been used to monitor the appearance of LYDMA during the virus-induced *in vitro* transformation of autologous B cells. Freshly prepared blood mononuclear cells from the relevant donors were first depleted of T cells by E-rosetting and one part of the T-depleted (TD) subpopulation was then exposed to a potent EB virus preparation, and the other left uninfected, before seeding in Linbro plates at 1.5×10^6 cells per 2-ml well. On at least two occasions during each experiment, the relevant effector T cells were tested in a 5-h chromium release assay against autologous TD cells from the uninfected cultures, against mixtures containing these TD cells and a 10, 25 or 50% proportion of cells from the autologous virus-transformed line, and against the autologous cell line itself. The results provided a calibration curve of effector T-cell activity, such as that shown in Fig. 1a for effectors

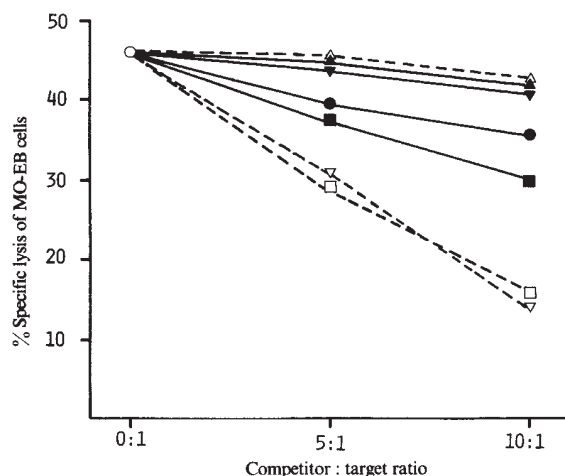


Fig. 2 Induction of LYDMA-positive cells in cultures of EB virus-infected TD cells from donor M.O. assayed by their capacity to serve as competitors of EB virus-specific cytotoxicity. Cultures of EB virus-infected TD cells from donor M.O. were collected and the cells stored in liquid nitrogen 1, 3, 4 and 7 days postinfection; in these preparations immunofluorescence staining showed here to be 1, 33, 53 and 74% EBNA-positive cells respectively. Effector T cells were prepared from donor S.G. (matched to donor M.O. at all four HLA-A and -B loci) and were tested for cytotoxicity against ⁵¹Cr-labelled target cells of the EB virus-transformed cell line from donor M.O. (MO-EB) either alone (○) or in the presence of the following unlabelled competitor cells: EB virus-infected MO-TD cells collected 1 (▲), 3 (▼), 4 (●) and 7 (■) days post-infection, prepared as above and thawed immediately before use; EB virus-transformed MO-EB cells (□); EB virus-transformed SG-EB cells (▽); TD cells from donor S.G. stimulated for 6 days in the presence of PWM and autologous mitomycin C-treated T cells as described elsewhere¹³ (Δ). When simultaneously tested against a range of target cells, the effector T cells from donor S.G. used above produced the following levels of specific lysis on (1) EB virus-transformed target cells: SG-EB (autologous), 42%; MO-EB (HLA-A and -B antigen-identical), 46%; HD-EB (HLA-A and -B antigen-unrelated), 3%; JR-EB (HLA-A and -B antigen-unrelated), 1%; (2) EB virus genome-negative target cells: PWM-stimulated SG-TD cells (autologous), -7%; PWM-stimulated MO-TD cells (HLA-A and -B antigen identical), -3%; K562, 1%; HSB-2, 7%.

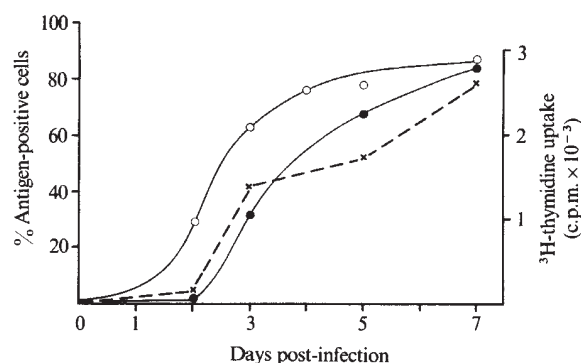


Fig. 3 Induction of EBNA-positive cells (○), LYDMA-positive cells (●) and DNA synthesis (×) in cultures of EB virus-infected TD cells from donor S.G. Assays for EBNA and LYDMA were as described in Fig. 1 legend. The effector T cells from donor S.G. used to detect LYDMA-positive cells in this experiment showed a pattern of EB virus-specific, HLA-A and -B antigen-restricted cytotoxicity similar to that described in Fig. 2 legend; the absolute level of autologous cell line lysis mediated by these effector T cells was 50–60% specific chromium release. Cellular DNA synthesis was measured by labelling the cultures with ³H-thymidine (TRA61, Amersham; specific activity 5 Ci mM⁻¹; concentration 0.5 μCi ml⁻¹ culture medium) for 18 h immediately before collection. Replicate 20-μl samples from each cell suspension were washed well with saline, followed by 10% ice-cold trichloroacetic acid and ice-cold methanol. The filters were dried and the remaining radioactivity measured by scintillation counting. The difference between the mean count obtained from virus-infected cultures and that obtained from parallel uninfected cultures was taken to indicate the level of virus-induced DNA synthesis.

from seropositive donor D.M., and confirmed that the assay was sufficiently sensitive to detect as few as 10% virus-transformed (LYDMA-positive) cells in a mixed target cell population.

Virus-infected TD cell cultures were collected on up to five occasions during the first week post-infection (including those days on which the above calibrations were performed), the cells counted and then used either to prepare slides for immunofluorescence staining to determine the proportion of EBNA-positive cells, or as targets for the relevant effector T cells in a chromium release assay (using the autologous cell line as a standard target on each occasion) to monitor the appearance of LYDMA-positive cells. All seven donors tested in this initial series of experiments gave results of the kind shown in Fig. 1b, again from donor D.M. Bright EBNA-positive staining, detectable in a small proportion of cells on day 1, was clearly present in 25% or more cells by day 2, rising steadily thereafter to levels of around 80% within 6–7 days. By comparison, virus-infected TD cells never suffered significant lysis in chromium release assays on days 1 or 2 post-infection but thereafter became increasingly sensitive such that about 80% of the cells were detectably LYDMA-positive by days 6–7.

The above chromium release data seems to give a true indication of the kinetics with which cells develop the T-lymphocyte-detected antigen (rather than merely reflecting some temporal change in the sensitivity of antigen-bearing cells to lysis *per se*) because a very similar pattern of results was obtained on three further occasions when the appearance of LYDMA-positive cells was monitored by testing virus-infected TD cell populations as unlabelled competitors of EB virus-specific cytotoxicity. Figure 2 shows the results from one such experiment using the EB virus-transformed cell line from donor M.O. as the source of target cells and a reactivated T-cell preparation from the HLA-A and -B antigen-identical donor S.G. as the source of effectors. When sequentially collected populations of EB virus-infected MO-TD cells were included as competitors in this assay, they did not begin to cause significant inhibition of lysis (that is, above the very low background effect of PWM-stimulated B-lymphoblast competitors) until day 4 post-infection even though some 33% EBNA-positive cells were already detectable by day 3. The discordance between the two antigen levels which was consistently observed on days 2–3 post-infection, and the rapid increase to equivalent values thereafter, strongly suggest that the appearance of EBNA in the nuclei of

Table 1 Development of EBNA and LYDMA in thymidine-blocked cultures

Cells in culture	Treatment	Day of collection	Cell count per ml	% EBNA-positive cells	% LYDMA-positive cells
SG-TD	EBV	5	0.35 × 10 ⁶	78	68
		7	10 ⁶	86	84
SG-TD	EBV + 10 mM Tdr	5	0.12 × 10 ⁶	70	81
		7	0.15 × 10 ⁶	76	88

Unlabelled thymidine (Tdr) was added to the cultures 24 h post-infection to a final concentration of 10 mM. The establishment of a block to cell proliferation, exerted at the first G₁-S boundary in thymidine-treated cultures, was confirmed by Hoechst 33342-staining analysis²⁴ of the cell populations shown above; when analysed in a fluorescence activated cell sorter, >95% thymidine-treated cells had pre-synthetic DNA values characteristic of the G₁ phase, whereas non-thymidine-treated cells showed a heterogeneous spread of DNA values across the cell cycle (data not shown).

virus-infected B cells precedes by about 24 h the expression of LYDMA in a recognizable form on the cell surface.

In three of the seven experiments of the initial series, both control and virus-infected TD cell cultures were labelled with ³H-thymidine for 18 h immediately before collection so as to assess the degree of virus-induced cell proliferation, in addition to the levels of viral antigen expression. Typical results, with cultures from donor S.G. (Fig. 3), show the induction of LYDMA and the initiation of cellular DNA synthesis to be coincident events occurring soon after the appearance of EBNA in virus-infected cells. In these same experiments, exposure of some virus-infected cultures to 10 mM unlabelled thymidine from day 1 onwards stopped the subsequent entry of cells into S phase, a block confirmed by Hoechst 33342-staining analysis of the cell populations in a fluorescence-activated cell sorter²⁴, but did not prevent a substantial majority of the infected cells going on to develop both EBNA and LYDMA (Table 1).

A close temporal relationship between the expression of EBNA, a DNA-binding protein thought to induce cell proliferation²⁵, and that of LYDMA, a cell-surface change which elicits T-cell-mediated attack^{15,18}, has important implications for the biology of the EB virus carrier state. Two facts, the remarkably strong long-lived T-cell surveillance which EB virus infection usually induces in its host⁶, and the obligatory appearance of LYDMA on cells at an early stage of the infectious cycle (Figs 1b, 2) coincident with, but not dependent on, the initiation of cell proliferation (Fig. 3, Table 1), are difficult to accommodate within that current view of EB virus persistence which postulates the long-term survival of individual virus-transformed B cells in lymphoid tissues. This raises the possibility of an alternative view, namely that the virus might persist by way of chronic, usually low grade, replication within the oropharynx²⁶⁻²⁸, involving some specialized cell and/or some immunologically privileged site, and that the apparent virus-carrier status of lymphoid tissues may be maintained only by the continual in-flow of newly infected B cells from that site coupled with their continual removal by immune T cells once the stage of LYDMA expression has been reached. Relaxation of T-cell surveillance in immunologically compromised individuals would of course allow EBNA/LYDMA-positive cells to begin proliferating and this may be a necessary first step in the pathogenesis of EB virus-associated malignant disease.

We thank Elizabeth Gough and Ann Green for technical help, and Dr B. Thomas and Mr K. Keeler for advice and assistance with the fluorescence-activated cell sorter analysis of Hoechst 33342-stained cell preparations. This work was supported by MRC grants G976/412 and G979/1233, and was assisted by the Bradbury Investment Company of Hong Kong. D.J.M. was supported by a D.B. Duncan Cancer Research Fellowship while on leave from the Queensland Institute of Medical Research, Herston, Brisbane, Australia.

Received 5 November 1980; accepted 17 March 1981.

- Epstein, M. A. & Achong, B. G. *Rev. Microbiol.* **31**, 421-445 (1977).
- Klein, G. *Cancer* **45**, 2486-2499 (1980).
- Henle, W., Diehl, V., Kohn, G., zur Hausen, H. & Henle, G. *Science* **157**, 1064-1065 (1967).

- Pope, J. H., Horne, M. K. & Scott, W. *Int. J. Cancer* **3**, 857-866 (1968).
- Nilsson, K., Klein, G., Henle, W. & Henle, G. *Int. J. Cancer* **8**, 443-450 (1971).
- Moss, D. J., Rickinson, A. B. & Pope, J. H. *Int. J. Cancer* **22**, 662-668 (1978).
- Rickinson, A. B., Moss, D. J. & Pope, J. H. *Int. J. Cancer* **23**, 610-617 (1979).
- Moss, D. J., Rickinson, A. B. & Pope, J. H. *Int. J. Cancer* **23**, 618-625 (1979).
- Rickinson, A. B., Wallace, L. E. & Epstein, M. A. *Nature* **283**, 865-867 (1980).
- Reedman, B. M. & Klein, G. *Int. J. Cancer* **11**, 499-520 (1973).
- Rickinson, A. B., Moss, D. J., Allen, D. J., Wallace, L. E., Rowe, M. & Epstein, M. A. *Int. J. Cancer* (in the press).
- Misko, I. S., Moss, D. J. & Pope, J. H. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4247-4250 (1980).
- Moss, D. J., Wallace, L. E., Rickinson, A. B. & Epstein, M. A. *Eur. J. Immun.* (in the press).
- Wallace, L. E., Moss, D. J., Rickinson, A. B., McMichael, A. J. & Epstein, M. A. *Eur. J. Immun.* (in the press).
- Rickinson, A. B., Moss, D. J., Wallace, L. E., Rowe, M., Misko, I. S., Epstein, M. A. & Pope, J. H. *Cancer Res.* (in the press).
- Svedmyr, E. & Jondal, M. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1622-1626 (1975).
- Royston, I., Sullivan, J. L., Periman, P. O. & Perlman, E. *New Engl. J. Med.* **293**, 1159-1163 (1975).
- Klein, E., Klein, G. & Levine, P. H. *Cancer Res.* **36**, 724-727 (1976).
- Ortaldo, J. R., Bonnard, G. D., Kind, P. D. & Herberman, R. B. *J. Immun.* **122**, 1489-1494 (1979).
- Sealey, J. K., Masucci, G., Poros, A., Klein, E. & Golub, S. H. *J. Immun.* **123**, 1303-1311 (1979).
- Viallat, J., Svedmyr, E., Steinitz, M. & Klein, G. *Cell. Immun.* **38**, 68-75 (1978).
- Jondal, M. & Targan, S. *J. exp. Med.* **147**, 1621-1636 (1978).
- Van Lambalgen, R., Farrant, J. & Bradley, B. A. *J. Immun. Meth.* **27**, 327-338 (1979).
- Lydon, M. J., Keeler, K. D. & Thomas, D. B. *J. cell. Physiol.* **102**, 175-181 (1980).
- Luka, J., Siegert, W. & Klein, G. *J. Virol.* **22**, 1-8 (1977).
- Gerber, P., Nonoyama, M., Lucas, S., Perlman, E. & Goldstein, L. I. *Lancet* **ii**, 988-989 (1972).
- Miller, G., Niederman, J. C. & Andrews, L.-L. *New Engl. J. Med.* **288**, 229-232 (1973).
- Morgan, D. G., Niederman, J. C., Miller, G., Smith, H. W. & Dowdall, J. M. *Lancet* **ii**, 1154-1157 (1979).
- Crawford, D. H., Rickinson, A. B., Finerty, S. & Epstein, M. A. *J. gen. Virol.* **38**, 449-460 (1978).

Inserted H-2 gene membrane products mediate immune response phenotype of antigen-presenting cell

Aya Jakobovits*, Ayala Frenkel†, Nathan Sharon* & Irun R. Cohen†

Departments of Biophysics* and Cell Biology†, The Weizmann Institute of Science, Rehovot, Israel

Soluble protein antigens do not stimulate helper or effector T lymphocytes unless they have first been processed and exposed to the T lymphocytes by appropriate antigen-presenting cells. This antigen presentation was found to be controlled by products of genes in the major histocompatibility complex (MHC), designated H-2 in the mouse. Efficient antigen presentation to T lymphocytes is obtained only when both presenting and responding cells share alleles within the I region of the MHC¹. Furthermore, immune response (Ir) genes in the MHC influence the choice of antigenic determinants recognized on the immunogenic molecule². To investigate the molecular mechanisms underlying Ir gene expression, we have inserted membrane components isolated from cells of one Ir genotype into recipient cells of a different Ir genotype, using membrane vesicles composed of donor plasma membranes and envelope glycoproteins from Sendai virus³. We now report that this 'membrane-engineering' enables genetically inappropriate antigen-presenting cells to communicate to T lymphocytes antigenic determinants of beef or sheep insulins, antigens under control of Ir genes in mice⁴. This demonstrates that structural products of MHC genes, expressed and active at the plasma membrane, function in the Ir phenotype of the response to insulin.

We immunized H-2 congenic mice, C3H · DiSn (H-2^k) or C3H · SW (H-2^b), against sheep or beef insulin, respectively. Although these two insulins differ by substitution of only one amino acid at position 9 in the A chain (sheep A9 Gly, beef A9 Ser), the products of H-2-linked Ir genes are sensitive to this structurally conservative difference. Mice of the H-2^b genotype respond well to beef insulin but only poorly to sheep insulin,

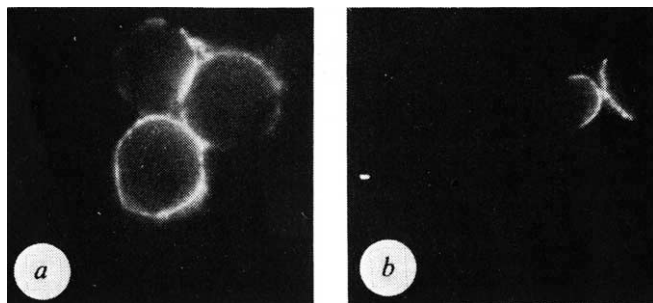


Fig. 1 H-2^b antigens inserted into membranes of H-2^k acceptor cells are capped by anti-H-2^b antiserum, detected by immunofluorescence. C3H · DiSn (H-2^k) spleen cells were fused with membrane vesicles containing C3H · SW (H-2^b) plasma membranes. Fusogenic membrane vesicles were prepared by mixing 1 mg protein of freshly solubilized plasma membranes with 1 mg protein of solubilized Sendai virus envelopes and dialysing the mixture for 96 h at 4 °C in the presence of Bio-beads SM-2 for efficient removal of the detergent³. The dialysed solution was centrifuged (100,000g, 3 h, 4 °C) to obtain reconstituted vesicles. Fusion procedure: 10⁷ C3H · DiSn (H-2^k) spleen cells were suspended in 1 ml of fusion solution (0.14 M NaCl, 20 mM Tris-HCl pH 7.4, 3 mM KCL, 0.8 mM MgSO₄) and incubated with 5 µg protein of reconstituted membrane vesicles for 60 min at 4 °C, with occasional shaking, during which the vesicles bound to the cells. To complete the process of fusion, the cells were collected by centrifugation, resuspended in 1 ml of fusion solution containing 5 mM Ca²⁺, incubated at 37 °C for 30 min and then washed twice with phosphate-buffered saline, pH 7.4. To detect H-2^b antigens on the surface membranes of the H-2^k cells (a), the cells were incubated for 30 min at 4 °C with anti-H-2^b antiserum (prepared by immunizing C3H · DiSn (H-2^k) mice with C3H · SW (H-2^b) spleen cells) followed by incubation with fluorescein isothiocyanate-labelled protein A for 30 min at 4 °C. To produce capping (b), the cells from a were incubated for additional 15 min at 37 °C (ref. 7). Negative results were obtained using control antiserum or H-2^k acceptor cells that had not been fused with membrane vesicles.

whereas H-2^k mice respond only to sheep insulin^{4,5}. We examined whether memory T lymphocytes from immunized mice would respond in a secondary proliferation assay *in vitro* to antigen-presenting spleen cells that had been pulsed with one of these insulins. It has been shown that insulin-pulsed spleen cells of an incompatible H-2 genotype will not present insulin to memory T lymphocytes¹, as T lymphocytes seem to recognize their specific antigen only in association with appropriate H-2 gene products⁶. Our approach was to insert membrane components isolated from cells of H-2^b or H-2^k responder genotype into the membrane of presenting cells of H-2 non-responder genotype, and to examine whether the inserted H-2 structural products could enable genetically inappropriate cells to present insulin to primed memory T lymphocytes.

To insert donor plasma membrane components into the membrane of recipient cells, we reconstituted isolated plasma membranes of spleen cells together with Sendai virus envelopes to obtain fusogenic membrane vesicles that could be integrated into the membranes of intact acceptor cells without killing them³. Isolation of purified plasma membranes and preparation of fusogenic membrane vesicles were carried out as previously described³. Antigen-presenting cells were fused with fusogenic membrane vesicles from H-2 congenic sources and the integration of the new H-2 antigens in the cell membrane was detected by binding specific anti-H-2 antiserum. As shown in Fig. 1, fusion of C3H · DiSn (H-2^k) cells with vesicles prepared from C3H · SW (H-2^b) spleen cell plasma membranes led to the appearance of H-2^b antigens on the surface of H-2^k cells. Capping⁷ of the inserted H-2 antigens indicated that these antigens were functionally integrated into the membranes of the

acceptor spleen cells. Moreover, it had been previously shown that inserted allogeneic H-2 antigens, but not unfused allogeneic vesicles, could also stimulate a primary mixed-lymphocyte reaction in syngeneic responder lymphocytes³.

Figure 2 shows the results of inserting H-2^k or H-2^b plasma membranes into H-2 syngeneic or allogeneic antigen-presenting spleen cells. In detecting the specific response to insulin, we had to take into account the fact that allogeneic H-2 antigen endogenous or inserted into spleen cells will stimulate T lymphocytes⁸. The net response to insulin was therefore the c.p.m. obtained when the antigen-presenting spleen cells were pulsed with insulin minus the c.p.m. generated in the absence of insulin pulsing—the stimulation due to allogeneic H-2 antigens was considered as background.

In the first experiment, lymph node lymphocytes were taken from C3H · DiSn (H-2^k) responder mice that had been immunized *in vivo* against sheep insulin. Antigen-presenting cells originated from C3H · SW (H-2^b) mice which are genetical low responders to sheep insulin. These H-2^b spleen cells, either untreated or fused with vesicles containing virus envelope components alone, or together with C3H · SW (H-2^b) membrane antigens, stimulated a response to their allogeneic H-2^b antigens, but did not stimulate a net response to insulin. However, insertion of C3H · DiSn (H-2^k) membrane components into the H-2^b antigen-presenting cells led to a strong net response to sheep insulin of 65,000 c.p.m. above the allogeneic background level measured in the absence of insulin. Similarly, fusion of C3H · SW (H-2^b) antigen-presenting cells with C3H/eB (H-2^k) membranes led to a strong net proliferative response to sheep insulin (62,000 c.p.m.).

The effect of membrane insertion on the ability of spleen cells to present beef insulin was also studied. The responding cells were lymph node cells taken from C3H · SW (H-2^b) responder mice that had been primed against beef insulin while antigen-presenting cells originated from C3H · SW (H-2^b) responder mice or from C3H · DiSn (H-2^k) nonresponder mice. C3H · SW (H-2^b) spleen cells presented beef insulin to syngeneic responding cells (net response 28,000 c.p.m.). The insertion of nonresponder allogeneic C3H · DiSn (H-2^k) membranes into the C3H · SW (H-2^b) spleen cells led to a significant allogeneic stimulation, but the net response to beef insulin was 19,000 c.p.m. about 30% less than that elicited by untreated C3H · SW (H-2^b) spleen cells. Antigen-presenting spleen cells from C3H · DiSn (H-2^k) mice produced an allogeneic response of about 40,000 c.p.m. with a negligible net response to beef insulin (7,000 c.p.m.). However, fusion of these H-2^k spleen cells with C3H · SW (H-2^b) membrane vesicles produced a strong net response to beef insulin of 62,000 c.p.m. above the allogeneic background stimulation. In contrast, insertion of vesicles containing virus envelope, either alone or with C3H · DiSn (H-2^k) membranes (not shown), into the C3H · DiSn (H-2^k) spleen cells did not improve the capacity of these cells to present beef insulin.

These results show that genetically inappropriate antigen-presenting spleen cells can be made to present insulin by inserting membrane components taken from H-2 genetically appropriate spleen cells. This was probably due to specific products of H-2 genes, because the donor and acceptor cells were congenic and differed only at their H-2 alleles. Presentation of insulin could not be attributed to any nonspecific allogeneic effects⁹ because, as reported by others¹, we observed that unmodified allogeneic spleen cells did not present antigens to primed memory T lymphocytes (Fig. 2 a, b), and insertion of membranes from nonresponder H-2^k genotype into H-2^b responder spleen cells led to a strong allogeneic response that not only did not enhance, but decreased the presentation of beef insulin by these cells to syngeneic H-2^b lymphocytes (Fig. 2a). In addition, insertion of viral envelope components alone or together with nonresponder H-2 membranes showed no effect on presentation of insulin. Thus, the ability of the inserted membrane components to change the presentation properties of the recipient spleen cells seemed to be specific to the H-2 alleles

Fig. 2 Inserted H-2 gene products mediate Ir phenotype of spleen cells presenting insulin to primed T lymphocytes. Mice of congenic strains C3H·DiSn (H-2^k) or C3H·SW (H-2^b) were immunized by injecting 5 µg of insulin emulsified in complete Freund's adjuvant into each hind footpad, as described elsewhere⁵. Nine days later, suspensions of lymphocytes from the draining popliteal lymph nodes were suspended in RPMI-1640 medium supplemented with 0.5% normal mouse serum, HEPES buffer (10×10^{-3} M, pH 7.3), non-essential amino acids (Bioblab, 1%), sodium pyruvate (1×10^{-3} M), L-glutamine (2×10^{-3} M), 2-mercaptoethanol (5×10^{-5} M) and antibiotics. 4×10^5 cells were seeded in U-shaped microtitre wells (Greiner) in 0.1 ml medium. Antigen-presenting spleen cells were untreated or fused with vesicles as described in Fig. 1 legend. Vesicles containing virus alone were prepared by reconstituting virus envelopes without the addition of plasma membranes. After fusion, the antigen-presenting cells (10^7 per ml RPMI-1640) were incubated with or without insulin ($100 \mu\text{g ml}^{-1}$) for 60 min at 37 °C, irradiated (1,000 rad) to inhibit their synthesis of DNA and washed twice in medium. The antigen-presenting cells were added (10^5 cells per 0.1 ml medium) to each well containing primed lymph node lymphocytes and the cells mixtures were incubated for 96 h in quadruplicate cultures. To measure the proliferative response, ³H-thymidine (1 µCi per well) was added to each well for the last 16 h of culture and incorporation into DNA was measured by collecting the cells and measuring the c.p.m. as described⁵. s.e.m. were all <10%. The net response to insulin was computed by subtracting the c.p.m. obtained in cultures without insulin-pulsed antigen-presenting spleen cells from the c.p.m. obtained in cultures with insulin-pulsed antigen-presenting spleen cells. The results of the groups showing net response to insulin of $\geq 19,000$ c.p.m. were highly significant by Student's *t*-test ($P < 0.001$). Non-fusogenic vesicles containing plasma membrane components without viral glycoproteins had no influence on the response to insulin (data not shown).

Responding T lymphocytes	Insulin-presenting spleen cells			Secondary proliferative response (c.p.m. $\times 10^{-3}$)				Net response to insulin (c.p.m. $\times 10^{-3}$)
	Origin of Spleen cells	Origin of inserted membranes	Pulsed with insulin	20	60	100	140	
a C3H·DiSn (H-2 ^k) primed to sheep insulin	C3H·SW (H-2 ^b)	None	SHEEP: +					0
		Virus alone	+ +					0
		C3H·SW (H-2 ^b)	+ +					0
		C3H·DiSn (H-2 ^k)	+ +					65
		C3H/eB (H-2 ^b)	+ +					62
b C3H·SW (H-2 ^b) primed to beef insulin	C3H·SW (H-2 ^b)	None	BEEF: +					28
		C3H·DiSn (H-2 ^k)	+ +					19
		C3H·DiSn (H-2 ^k)	+ +					7
		Virus alone	+ +					4
		C3H·SW (H-2 ^b)	+ +					62

and Ir genes that distinguished the congenic strains of mice used.

As well as modifying the Ir phenotype of the insulin-presenting cells insertion of syngeneic H-2 membrane products into allogeneic cells augmented the background allogeneic stimulation (Fig. 2). The recognition of allogeneic antigens may have been enhanced by association with self-H-2 products in the same membrane⁶.

It has been suggested that presentation of antigens involves two distinct processes coded by H-2 genes: the communication of signals between cells mediated by 'cell-interaction structures'¹⁰, and the selection of specific antigenic determinants on the immunogenic molecule². Our results indicate that structural products of H-2 genes located at the plasma membrane of the presenting cells are sufficient to carry out these functions. A strategy of membrane reconstruction using isolated and purified H-2 gene products will make it possible to study the molecular mechanisms responsible for these cellular functions.

This study was supported in part by grant AM26766 from the National Institute of Arthritis, Metabolism and Digestive Diseases. Dr P. De Baetselier donated the fluorescein isothiocyanate-labelled Protein A.

Received 8 December 1980; accepted 17 March 1981.

- Yano, A., Schwartz, R. H. & Paul, W. E. *J. exp. Med.* **146**, 828–843 (1977).
- Rosenthal, A. S., Barcinski, M. A. & Blake, J. T. *Nature* **267**, 156–158 (1977).
- Prujansky-Jakovits, A., Volsky, D. J., Loyter, A. & Sharon, N. *Proc. nat. Acad. Sci. U.S.A.* **77**, 7247–7251 (1980).
- Keck, K. *Eur. J. Immun.* **7**, 811–816 (1977).
- Cohen, I. R. & Talmon, J. *Eur. J. Immun.* **10**, 284–289 (1980).
- Snell, G. D. *Immun. Rev.* **38**, 3–69 (1978).
- Taylor, R. B., Duffus, W. P. H., Raff, N. C. & De Petris, S. *Nature new Biol.* **233**, 225–229 (1971).
- Bach, F. H., Widmer, M. B., Bach, M. L. & Klein, J. *J. exp. Med.* **136**, 1430–1444 (1972).
- Osborne, D. P. Jr & Katz, D. H. *J. exp. Med.* **138**, 825–838 (1973).
- Katz, D. H. & Benacerraf, B. *Transplant. Rev.* **22**, 175–195 (1975).

Differences between germ-line and rearranged immunoglobulin V_κ coding sequences suggest a localized mutation mechanism

Michael Pech, Josef Höcht, Hannelore Schnell & Hans G. Zachau

Institut für Physiologische Chemie, Physikalische Biochemie und Zellbiologie der Universität München, 8000 München 2, FRG

Several mechanisms contribute to the generation of an antibody repertoire (for reviews see refs 1–3) which has been estimated to exceed 10^6 different types of immunoglobulin molecules⁴. It is clear that the germ line contains a fair number of V-gene segments^{2,5–7}. The V-gene segments are joined by somatic recombination to J- and D-gene segments^{8–11} and there is, of course, combinatorial association of heavy and light chains. It has been suggested that the germ-line repertoire of gene segments and their combinatorial joining are the principal sources of antibody diversity¹². In fact, two rearranged V_κ-gene segments have been found to be identical in sequence to their germ-line counterparts^{12,13}, showing that no somatic point mutations had occurred. In the genes for a λ₁ light chain¹⁴ and for two heavy chains^{10,15}, on the other hand, the rearranged V-gene segments differed in several positions from their germ-line equivalents; this is explained by somatic point mutations. We now report the finding of six single-base pair differences each between two rearranged V_κ-gene segments and their presumptive germ-line equivalents and the absence of any such differences in the adjacent non-coding regions; this suggests the existence of localized mutation mechanism(s) in this V_κ gene system.

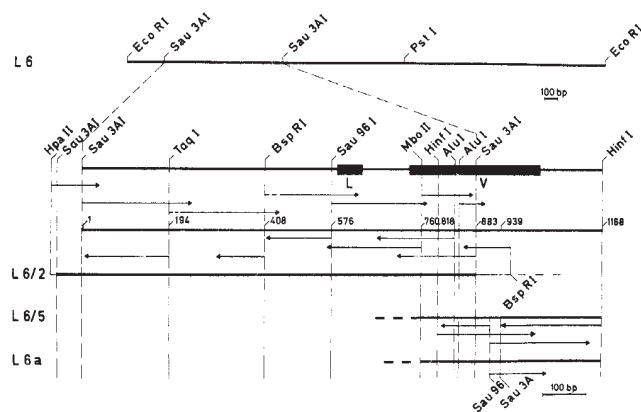


Fig. 1 Partial restriction maps and sequencing strategies of L6 and the subclones L6/2 and L6/5. The subclone L6/2 was constructed by ligation of a *Sau3AI* digest of L6 into the *Bam*HI site of pBR322. For the construction of subclone L6/5 the leftward *EcoRI*-*PstI* fragment of L6 was ligated with an *EcoRI*-*PstI* digest of pBR322. In L6/2 the *Sau3AI* fragment 884-939 was ligated onto the left end of fragment 1-883. The restriction maps of subclone L6/5 and the second independently isolated λ gtWES clone L6a (bottom line) (see text) are shown only in the regions relevant for the sequencing strategy. Thick lines designate mouse DNA, thin lines pBR322 DNA. Digestions, isolation of DNA fragments and further steps of sample preparation were as in ref. 24 with modifications according to ref. 17.

The variable regions of two rearranged κ light-chain genes from the same mouse tumour, myeloma T (ref. 16), have recently been sequenced¹⁷. The two gene segments which had been generated by joining to the J5 and J2 gene segments (designation as in ref. 8) were called V-T1 and V-T2, respectively¹⁶. The V-T1 gene segment is expressed (ref. 17 and F. Falkner, unpublished) whereas V-T2 is non-functional because of a one-nucleotide deletion at the V-J junction. Presumptive germ-line counterparts were cloned from liver DNA of BALB/c mice^{7,8}, which in blotting experiments of digests was identical to embryonic DNA¹⁶. Column chromatographic fractions of *EcoRI* digests of liver DNA were screened by hybridization with cloned fragments containing V-T1 and V-T2 (Fig. 1 in ref. 7). Five fragments hybridizing with V-T1 were cloned, four of which can be excluded as germ-line equivalents of V-T1 on the basis of restriction analyses^{7,18}. Other fragments were excluded because of their weak hybridization signals. The fragment designated L6, however, which gave rise to the strongest hybridization signal, had identical restriction patterns in the V-gene region. The same pattern was found in three independently isolated clones. V-L6 is therefore the presumptive germ-line equivalent of V-T1 (ref. 7). Screening with a V-T2-containing clone revealed three weakly hybridizing fragments which differed from V-T2 in restriction patterns⁷. The only strongly hybridizing fragment (fractions 211-219 in Fig. 1 of ref. 7), designated L7, was cloned¹⁸. V-L7 is the presumptive germ-line equivalent of V-T2. Also, for V-L7 three independently isolated clones had identical restriction patterns. Subclones of L6 and L7 were constructed (Figs 1, 3) and the pertinent regions sequenced (Figs 2, 4).

Fig. 2 Nucleotide sequence of the V-gene region of L6. The sequence extends at the 5' end across the *Sau3AI* site (see Fig. 1) but numbering was made consistent with Fig. 2 of ref. 17. Differing nucleotides or amino acids in V-T1 are indicated as superscripts or subscripts, respectively. The 3'-flanking region starts at position 1,028.

```

CAATGTTATATGTGCAACTCCTTTGTATTTCTCTCTTTGAACTGTAGTACAAAGCAGTAGAACCTTATGTTTAAATTCGGTGAAAATGTTTAT 100
GATAAATCTCCTTTTAAATTCACAAATGCCTACTCCAAACACTTCACTAGGGAGGAAATGCTATTATTTAGACACCAAGAAATGATCTTATCCAGAGT 200
AGTTAAAGATATAATAGTTTCTTTCTGTTGATATTTCCCTGACAGCAGGGGATCTTCTATATACCTTAGATATCTTGAATATATTATTCAAAT 300
AGAGACTACCTACAATAGAAATGACATCTATAATTATAAATCTTTGAAAGAAATTCATTTTCTGGAGATATTCAGTGGTTTGACCTCAGTCATCTGGG 400
GACACTGGCCTTGAAGTTAAGTCTGACATGAGCAAGCTTTGACCTGTGTCCAGCAATAAAGTGGTCCCAATGATTTCATGCTCTCTACCTCAGTGCCT 500
GGGGACTCTTCATATACCCGTCACACATGTACGGTACCATTTGCTATTCGACGACAGGACTCAGCATGGACATGAGGACCCCTGCTCAGTTCTTGGAAATC 600
TGTGTTGCTCTGGTTTCCAGGTAATAAGTAACTAAATGGGAATGTCAGTGTGATTAGTGTGATTGGCATTGGGAGATTTATCTTTTATGATGCTTACC 700
LeuLeuLeuTrpPhePro
TATGTAGATACTCATTTATGCTCCATTTCTGATATCAATGTACATCAAGTACCCAGTCTCCATCTTCCATGTATGATCTCTAGGAGAGAGAGTCT 800
GlyIleLeuCysAspIleLeuMetThrGlnSerProSerProSerMetTyrAlaSerLeuGlyGluArgVal
ACTATCTTGAAGGCGAGTCAGGACATTAATAGCTATTTAAGCTGGTTCCAGCAGAAACCCAGGAAATCTCCTAAGACCTGTATCTATCTGCAACCA 900
ThrIleThrCysLeuAlaSerGlnAspIleAsnSerTyrLeuSerTrpPheGlnGlnLeuProGlyLeuSerProLeuThrLeuIleTyrArgAlaAsnAla
GATTGGATGATGGGTCCTCATCAAGTTTCAGTGGCAGTGGATCTGGGCAAGATTATCTCTCACCATCAGCAGCCTGGAGTATGAAGATATGGGAATTTA 1000
rLeuValAspGlyValProSerArgPheSerGlySerGlySerGlyGlnAspThrSerLeuThrIleSerSerLeuGlyIleTyrGluAspMetGlyIleTyr
TATTGTCTCAGATATGATGAGTTTCTCCACAGTGAGACAAGTCATAACATTAACCCCAAGGAGAGAGAGGCTAGGCTGCTCAGCTGTT 1100
rTyrCysLeuGlnTyrAspGluPhePro
CCTCCTAATGTGTACCCACTGGTATTGTTCTAGATGTCACAGAACTTTAAGAAAGCTGTCACAGAGT

```

The 5'-flanking region of V-L6 as far as it was sequenced, the leader sequence, the small intron and the first 21 codons of the light-chain gene were found to be identical to the corresponding regions of V-T1 (1-806 in Fig. 2 and Fig. 2 of ref. 17). Further downstream, six differences were found, four of which also coded for different amino acids (Fig. 2). Three of the amino acid differences are located in framework regions and one occurs in the first hypervariable region¹⁹. Sequence comparisons between V-T1 (ref. 17) and J5 (refs 8,9) show that V-L6 terminates within the Pro codon 95 (position 1,027, Figs 2, 5a). In the 3'-flanking region the hepta- and decanucleotide sequences which are complementary to sequences upstream of the J-gene segments^{8,9} are found in the expected positions (1,031-1,037; 1,049-1,058).

The 5'-flanking region of V-L7, the leader sequence, the small intron and the first 49 codons of the light-chain gene are identical to the corresponding regions of V-T2 (1-466 in Fig. 4 and 755-1,220 in Fig. 5 of ref. 17). According to restriction patterns the homology extends further upstream^{16,18} (at least 6.2 kilobases from the leftward *Sau3AI* site in Fig. 3). V-L7 itself differs from V-T2 in 6 base pairs (bp) (Fig. 4), resulting in two different amino acid assignments in the framework region and three in the hypervariable regions. The conserved hepta- and decanucleotide boxes on the 3' side are also found in V-L7 in the expected positions; therefore, the aberrant joining found in V-T2 (ref. 17) cannot be explained by unusual features of these boxes. The joining point and, with that, the end of the coding region of V-L7, cannot be determined exactly because three different recombination events between CCAACCACA of V-L7 (positions 602-610) and GTGTACACG of J2 would lead to the same recombined sequence CCAACACGT in V-T2 (Fig. 5b). The observed reading frameshift in the J-gene segment of V-T2 (ref. 17) can thus be explained by deletion of T, A or C of the J2 sequence.

Although L6/T1 and L7/T2 cannot be assigned to a known subgroup or isotype, an amino acid sequence derived from the V-L6 sequence is identical to that expressed in the tumour RPC-5 which has been determined up to position 38 (S. Rudikoff, personal communication). V-L6 may therefore be the germ-line gene from which the V-gene segment of RPC-5 was formed without mutations within the first 38 codons, whereas on formation of V-T1 two such mutations were introduced in this region.

V-L6 and V-L7 have been called here the presumptive germ-line equivalents of V-T1 and V-T2, respectively. Are they really the germ-line equivalents? We cannot exclude the theoretical possibility that pseudoallelism²⁰ is the reason for the sequence differences; one would have to assume that the 'true' germ-line equivalents of V-T1 and V-T2 with the respective identical sequences were deleted from the germ line of our BALB/c mice after myeloma T has been induced, and the hybridization experiments with the V-T1 and V-T2 probes therefore show V-L6 and V-L7, respectively, to be the most closely related genes. On the basis of blot hybridization (M.P., unpublished results) it is unlikely (but cannot be excluded) that myeloma T, although it is propagated on BALB/c mice, is derived from another mouse strain which has undergone mutations in the

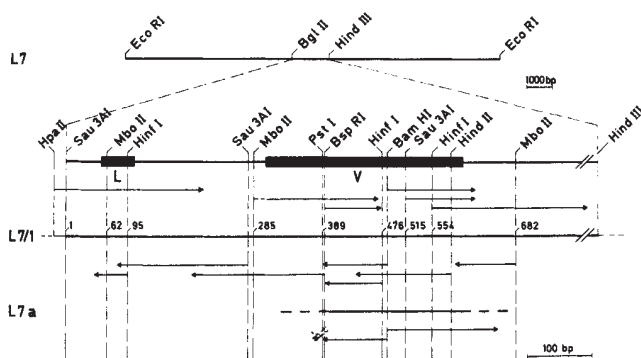


Fig. 3 Partial restriction maps and sequencing strategies of L7 and subclone L7/1. The subclone L7/1 was constructed by ligation of a *Bgl*II/*Hind*III digest with a *Bam*HI/*Hind*III digest of pBR322 DNA. The restriction map of the second independently isolated λ Charon 4A clone L7a (bottom line) is only shown for the region sequenced. Experimental conditions and presentation of the data are as in Fig. 1.

Fig. 4 Nucleotide sequence of the V-gene region of L7. Nucleotide positions 1–604/606 correlate to nucleotide positions 755–1,358/1,360 of V-T2 in Fig. 5 of ref. 17. Two alternative sites for the L-V gene splicing are indicated: splicing the leader codon 17 with the V-gene codon –1 would produce Gly 17, which is considered an invariant amino acid in light-chain leader sequences (see also ref. 17), while splicing with the codon –4 would produce an Ala 17 but conform better to the splice site consensus sequence and to regularities in leader length. The 3'-flanking region starts at 605/607 (see text). Differing nucleotides or amino acids are indicated as in Fig. 2.

```

GATCTTTCTCAGGGATGTGTCATGGTCCACAACTCAGGGAAGTTGAAGATGGTATCCACCTCAGTTCCTTTGCTTTTCTGGATTC
100
AGGATGACTGTTTGGGTGGCAAAAAGTGAGATGTTATTTAAATACAAAATTTCTGCTTTATTTGGAACCAATGTCACATGGGAATGACTTT
200
CAGTTTAAAGAAATGATACAATAAAAGTCATTATTTTCTAAGTGTGTTAGAAAGTACATTCATGATGTTGTGATCTACTAATGTATCTCTAT
300
TTTTCCAGCTCCAGAGGACATCTGTCTGACTCAGTCTCCAGCCATCCTGCTGTGAGTCCAGGAGAAAGAGTCAGTTCTCTCTGAGGGCCAGTCAG
400
AlaSerArgGlyAspIleLeuLeuThrGlnSerProAlaIleLeuSerValSerProGlyGluArgValSerPheSerCysArgAlaSerGln
AGCATTGGCACAAGCATACAGTGTATGCAAGAACAATGGTCTCCAGGCTCTCATAAAGTATCTCTGAGTCTATCTCTGGATCCCTTCCA
500
SerIleGlyThrSerIleHisTrpTyrGlnGlnArgThrAsnGlySerProArgLeuLeuIleLeuValAlaSerGluSerIleSerGlyIleProSerA
GGTTTAGTGCCAGTGGATCAGGACAGATTCTTACATGACATCAACAGTGTGGAGTCTGAAGATATCCAGATATTACTGTCAACAAGTAATAGCTG
600
roPheSerGlySerGlySerGlyThrAspPheThrLeuSerIleAsnSerValGluSerGluAlaAspTyrTyrCysGlnGlnSerAsnSerTr
GCCAACCCAGTGTGATGACACCATAGCAAAATCAGGAGAAAGCAGAGGCTGGTGTCTTATGCTATTTCTGGTGGTGTCTCAAACTCTTCAGAGTG
700
PPro
(95)
TTTCACAGATGCAGCCATGTGTGATGTTTACAAAATATGTTGAATGAAGAAGCCCTAAAGTCTGCTGACTGTAACTCTTTCGATTCTA

```

germ-line genes corresponding to T1 and T2. However, the simplest interpretation of our blot hybridizations and restriction mapping is the possibility that we are dealing here with the rearranged and non-rearranged counterparts.

Sequence differences have also been explained by residual allele heterogeneity or polymorphism among individual BALB/c mice^{10,14}. However, the finding of six sequence differences each within 203 and 131 bp, respectively, and of no differences in the large adjacent regions or in the *C_κ* gene region²¹, argue against polymorphism as the main cause. For the same reason it is unlikely that mutations as they were observed in aged spinner cultures of MOPC-21 (ref. 22) and as they may also occur during tumour propagation contribute to the differences between our germ-line and rearranged sequences.

The relevant regions of V-L6 and V-L7 were sequenced in two independently isolated clones each (Figs 1, 3). No sequence differences were detected. Also, the V-T2 sequence was based on three independent clones¹⁷. This argues strongly against the supposition that the observed sequence differences arose from cloning artefacts or from fortuitously picking single clones from families of fragments with identical restriction patterns but slightly differing sequences, although such possibilities may never be completely excluded.

The differences between the germ-line and the rearranged *V_κ* genes could be generated by a somatic mutation mechanism acting preferentially on the coding sequences (perhaps concomitantly with the V-J rearrangement) or by the absence in those sequences of a surveillance process operating elsewhere. It is clear that such a mechanism is not acting on all *V_κ* genes because Leder's group found the rearranged *V_κ* gene sequences of MOPC-41 (ref. 12) and MOPC-173B (ref. 13) to be identical to germ-line sequences. The results reported here, on the other hand, add to the known cases of somatic mutations investigated at the DNA^{10,14,15} and protein²³ level, a particularly clear case which, because of the clustering of several sequence differences

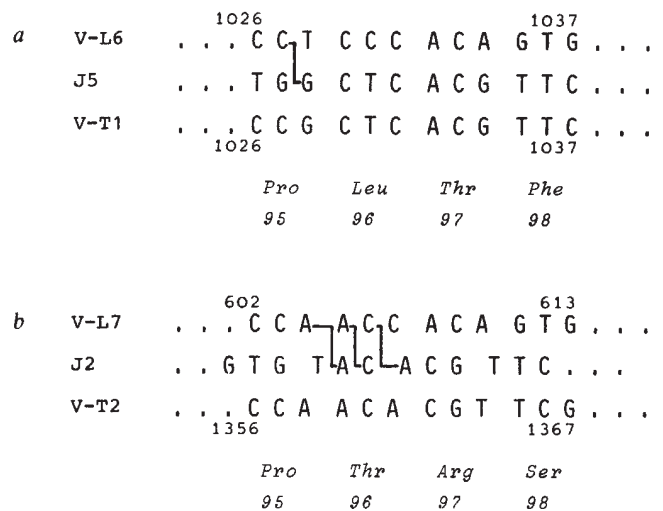


Fig. 5 Joining of V-L6 to J5 at position 1,027 (a) and V-L7 to J2 (b). In the latter case, joining at positions 604, 605 or 606 leads to the same (non-functional) sequence (see text). Numbering in V-L6 and V-L7 is as in Figs 2 and 4, in V-T1 and V-T2 as in ref. 17; the J2 and J5 sequences are according to refs 8,9.

in a small region and the absence of any differences in fairly large adjoining regions, argues in favour of a localized mutation mechanism. Further work must show to what extent such a mutation mechanism contributes to the generation of a functional antibody repertoire.

We thank R. Gerl for assistance, P. S. Neumaier for the computer programs, and particularly S. Rudikoff for communicating his unpublished sequence data. *Bsp*RI, *Hpa*II, *Mbo*II, *Sau*3AI and *Taq*I were from H. Feldmann, W. Gebhard and R. E. Streeck, and *Alu*I and *Bam*HI from Boehringer Mannheim. The work was supported by Bundesministerium für Forschung und Technologie.

Received 5 March; accepted 26 May 1981.

1. Tonegawa, S., Brack, C., Hozumi, N., Matthysens, G. & Schuller, R. *Immun. Rev.* **36**, 73–94 (1977).
2. Seidman, J. G., Leder, A., Nau, M., Norman, B. & Leder, P. *Science* **202**, 11–17 (1978).
3. Robertson, M. *Nature* **290**, 625–627 (1981).
4. Williamson, A. R. *Rev. Biochem.* **45**, 467–500 (1976).
5. Potter, M. *Adv. Immun.* **25**, 141–211 (1977).
6. Rabbitts, T. H., Matthysens, G. & Hamlyn, P. H. *Nature* **284**, 238–243 (1980).
7. Steinmetz, M., Höcht, J., Schnell, H., Gebhard, W. & Zachau, H. G. *Nucleic Acids Res.* **8**, 1721–1729 (1980).
8. Sakano, H., Hüppi, K., Heinrich, G. & Tonegawa, S. *Nature* **280**, 288–294 (1979).
9. Max, E. E., Seidman, J. G. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3450–3454 (1979).
10. Early, P., Huang, H., Davis, M., Calame, K. & Hood, L. *Cell* **19**, 981–992 (1980).
11. Rudikoff, S., Rao, D. N., Glaudemans, C. P. J. & Potter, M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4270–4274 (1980).
12. Seidman, J. G., Max, E. E. & Leder, P. *Nature* **280**, 370–375 (1979).
13. Max, E. E., Seidman, J. G., Miller, H. & Leder, P. *Cell* **21**, 793–799 (1980).
14. Bernard, O., Hozumi, N. & Tonegawa, S. *Cell* **15**, 1133–1144 (1978).
15. Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. & Tonegawa, S. *Nature* **286**, 676–683 (1980).
16. Steinmetz, M. & Zachau, H. G. *Nucleic Acids Res.* **8**, 1693–1707 (1980).
17. Altenburger, W., Steinmetz, M. & Zachau, H. G. *Nature* **287**, 603–607 (1980).
18. Höcht, J. thesis, Univ. München (1980).
19. Kabat, E. A., Wu, T. T. & Bilofsky, H. *NIH Publ.* 80-2008 (1979).
20. Hilschmann, N. *et al. Naturwissenschaften* **65**, 616–639 (1978).
21. Altenburger, W., Neumaier, P. S., Steinmetz, M. & Zachau, H. G. *Nucleic Acids Res.* **9**, 971–981 (1981).
22. Adetugbo, K., Milstein, C. & Secher, D. S. *Nature* **265**, 299–304 (1977).
23. Gearhart, P. J., Johnson, N. D., Douglas, R. & Hood, L. *Nature* **291**, 29–34 (1981).
24. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499–560 (1980).

Bleomycin-resistant DNA synthesis in ataxia telangiectasia cells

Philippe Cramer* & Robert B. Painter†

Laboratory of Radiobiology, University of California, San Francisco, California 94143, USA

Ataxia telangiectasia (AT) is a human autosomal recessive disease that exhibits (in addition to the symptoms implied by its name) immune deficiencies, greatly increased risk of lymphoreticular cancer and unusual sensitivity to ionizing radiation (see ref. 1 for review). Although cells from AT patients are abnormally sensitive to killing by X-ray irradiation, DNA synthesis in AT cells is resistant to such treatment^{2,3}. This resistance is primarily due to less inhibition of replicon initiation in AT cells than in normal cells, but there is also less inhibition of chain elongation in AT cells³. We now report that DNA synthesis in AT cells is also resistant to DNA damage induced by bleomycin but not to damage induced by *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (MNNG) or actinomycin D. Of these agents, only X rays and bleomycin directly cause DNA strand breaks, indicating that AT cells are defective in the system that in normal cells prolongs the time during which strand breaks can be repaired, replicon initiation inhibited and the chromatin structure returned to its normal state.

Normal human fibroblasts (HS27) and AT3BI cells were exposed for 30 min to various concentrations of bleomycin and their rates of DNA synthesis compared 30 min later (Fig. 1). The dose response for inhibition of DNA synthesis in normal cells showed a steep decline at low doses and a shallow decline at high doses, as previously reported⁴. It is generally assumed that the steep component is primarily due to inhibition of replicon initiation^{5,7} and the shallow one to blocked chain elongation in individual replicons⁸. The dose response for AT3BI cells, which are classified as excision deficient¹, showed no decrease in rate of synthesis at concentrations up to 75 $\mu\text{g ml}^{-1}$; at higher concentrations DNA synthesis was inhibited, but to a much lesser extent than in normal cells. In AT5BI cells, which are classified as excision proficient⁴, almost identical results were obtained (data not shown). These responses are similar to those observed with X-ray irradiation^{1,2}.

Sedimentation analysis of DNA on alkaline sucrose gradients confirmed the interpretation that bleomycin-induced DNA damage that inhibited replicon initiation in normal cells failed to do so in AT cells (Fig. 2). The major difference between the profiles of newly synthesized DNA from bleomycin-treated AT3BI cells and normal human fibroblasts is that, for the latter at 30 min after bleomycin treatment, there was a much larger reduction of radioactivity in low-molecular-weight regions (fractions 4–10) compared with untreated cells. This radioactivity reflects the part of DNA synthesis due to replicon initiation^{7–9}. (That inhibition of replicon initiation is the principle basis for the inhibition of DNA synthesis induced by low doses of X rays (<1,000 rad) has been confirmed by autoradiography¹⁰ and pH step alkaline elution⁹.) Chain elongation was also more resistant to bleomycin in AT3BI than in normal cells. In treated and untreated AT cells the high-molecular-weight regions of the profiles (fractions 13–24), which reflect chain elongation and joining of newly formed DNA^{7–9}, were almost superimposable, whereas the same regions of the profiles of DNA from treated normal cells show a reduction of radioactivity. These results cannot be explained on the basis of reduced entry of bleomycin into AT cells, because Taylor *et al.*¹¹ have shown that equal concentrations of bleomycin in the medium induce more chromosomal aberrations in AT

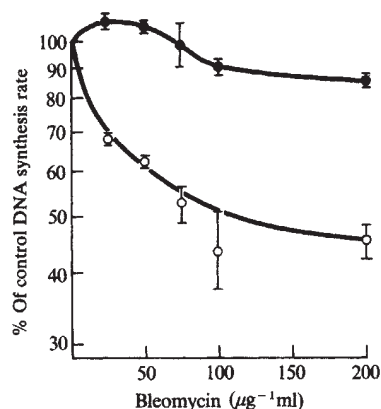


Fig. 1 Effect of bleomycin on rate of DNA synthesis in normal diploid HS27 (○) and AT3BI (●) cells. Cells grown in Eagle's minimal essential medium supplemented with 15% fetal bovine serum were inoculated into 35-mm Petri dishes at $\sim 5 \times 10^4$ cells per dish. After 1 day of growth the medium on all cultures was replaced with medium containing 0.01 $\mu\text{Ci ml}^{-1}$ of ^{14}C -thymidine (50 mCi mmol^{-1}) and incubation was continued for 24 h. The medium was removed and two dishes containing AT cells and two containing HS27 cells were incubated for 30 min in medium containing various concentrations of bleomycin, after which this medium was replaced with nonradioactive medium and the cells returned to the incubator for 30 min. The medium was then removed and all cultures were incubated with medium containing 20 $\mu\text{Ci ml}^{-1}$ of ^3H -thymidine (50 Ci mmol^{-1}) for 10 min; the cells were washed with ice-cold SSC (0.15 M sodium chloride, 0.015 M sodium citrate) and scraped into SSC. Ice-cold 4% perchloric acid was added and the cell suspension filtered through GF/C (Whatman) filters, which were rinsed sequentially with 4% perchloric acid, 70% alcohol and 100% alcohol, dried, and assayed for radioactivity in a liquid scintillation spectrometer. The resulting $^3\text{H}/^{14}\text{C}$ ratios were measures of the specific activity of DNA and therefore of the rate of DNA synthesis. Error bars indicate the range of duplicate samples.

cells than in normal cells. Moreover, preliminary data (not shown) indicate that equal numbers of single-strand DNA breaks are induced by bleomycin in AT and normal cells. Thus, AT cells are resistant to the bleomycin-induced damage that inhibits replicon initiation and blocks chain elongation in normal human cells.

As has been suggested for X-ray irradiation^{2,3}, the reason for the observed differences in DNA replication after bleomycin treatment may be that AT cells lack a factor or process that in normal cells intervenes to delay replication of damaged regions of DNA, or that AT cells have an altered chromatin structure that does not respond to bleomycin-induced damage. Previous studies have suggested that, in the chromatin of normal cells, DNA strand breaks induce a conformational change that inhibits initiation of DNA synthesis throughout a replicon cluster¹². If this is true, AT chromatin must not experience this conformational change after damage induced by bleomycin.

To determine whether the reported sensitivities of AT cells to killing by MNNG and actinomycin D^{13,14} are correlated with those of bleomycin, we measured the inhibition of DNA synthesis induced by these two agents (Fig. 3, Table 1). In neither case was there a significant difference between the two cell types in the effects of these agents on DNA synthesis.

Thus we suggest that high yields of DNA strand breaks are directly produced by X rays and bleomycin¹⁵ and are responsible for the inhibition of initiation of replicon clusters in normal cells. Actinomycin D and MNNG do not directly produce many strand breaks and therefore do not affect DNA synthesis at the level of the cluster, but instead induce lesions (methylated bases and intercalated molecules) that inhibit DNA replication primarily at the level of the individual replicon in both normal and AT cells. However, some of the 'X-ray-like' damage is induced by these agents. In AT cells, the defective damage-recognition system fails to respond to this damage and allows DNA

* Present address: Institut de Recherche d'Hématologie et d'Oncologie, Département de Pédiatrie, Clinique des Maladies du Sang, Hôpital Saint-Louis, 2, Place du Docteur Fournier, 75010 Paris, France.

† To whom correspondence should be addressed.

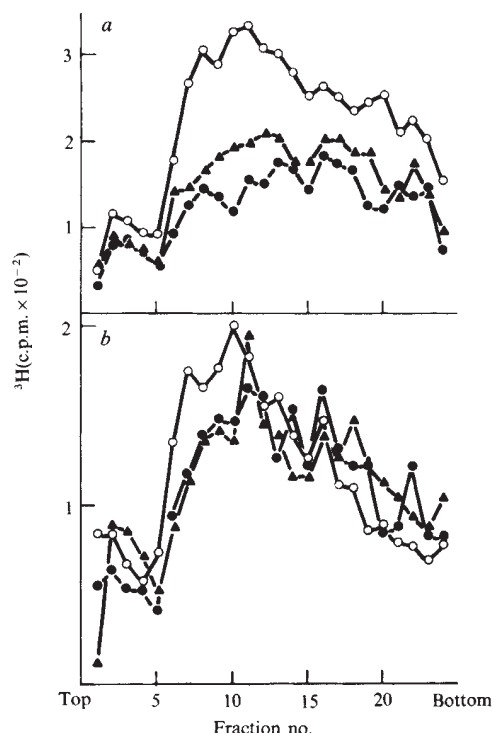


Fig. 2 Alkaline sucrose gradient profiles of newly formed DNA from normal diploid HS27 (a) and AT3BI (b) cells 30 min after treatment with bleomycin. ○, Untreated cells; ▲, 25 µg ml⁻¹ bleomycin; ●, 50 µg ml⁻¹ bleomycin. After the 30-min incubation in drug-containing medium, the cells were incubated in drug-free medium for 30 min and then in medium containing ³H-thymidine for 10 min. The cells were immediately cooled to ice temperatures, suspended in ice-cold buffer, irradiated with 1,000 rad to aid untangling of strands and lysed directly on top of 5–20% alkaline sucrose gradients. Centrifugation for 3 h at 27,000 r.p.m. in a Beckman SW-27 rotor was followed by collection of fractions, counting of radioactivity and analysis as described previously⁷.

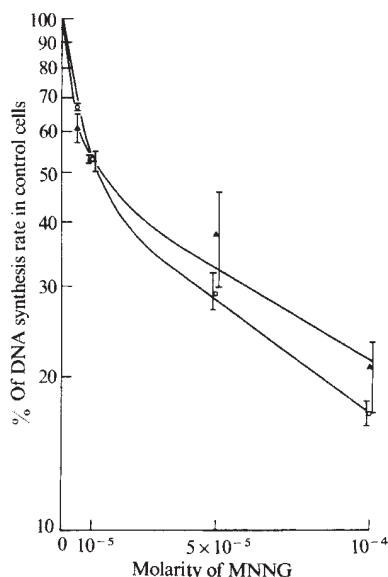


Fig. 3 Effect of MNNG on rate of DNA synthesis in normal diploid HS27 (○) and AT3BI (▲). The procedure is the same as that used with bleomycin (see Fig. 1 legend). Error bars indicate the range of duplicate samples.

replication to occur on damaged templates. This defective system also allows cells to progress from G₂ into mitosis before pre-aberrational damage is completely repaired, resulting in the reduced mitotic delay¹⁶ and increased frequencies of chromosomal aberrations¹⁷.

This hypothesis would not seem to explain the enhanced

Table 1 Inhibition of DNA synthesis by actinomycin D in AT and normal cells

Concentration of actinomycin D (M)	DNA synthesis (% of control) Normal cells	AT cells
5.0 × 10 ⁻⁷	76 ± 1.0	76 ± 2.0
7.5 × 10 ⁻⁷	60 ± 5.0	72 ± 2.0
1.0 × 10 ⁻⁶	47 ± 2.0	54 ± 8.0
2.0 × 10 ⁻⁶	40 ± 1.0	21 ± 5.0
5.0 × 10 ⁻⁶	11 ± 0.5	9 ± 1.0

The procedure was as used with bleomycin (see Fig. 1).

frequencies of chromosomal aberrations observed after irradiation of AT lymphocytes in G₀ (ref. 17), that is, before stimulation by a mitogen. However, Saha and Tolmach¹⁸ showed that HeLa DNA synthesis was inhibited on entry into S phase even when the cells had been irradiated several hours earlier while in G₁. If this phenomenon is due to a delay during which pre-aberrational lesions are repaired, the defective damage-recognition system in AT cells may cause them to progress precociously through S phase before repair is complete and thus allow the eventual expression of the lesion in mitosis.

This work was supported by the US Department of Energy and by a grant from the Fondation de l'Industrie Pharmaceutique pour la Recherche, Paris.

Received 31 October 1980; accepted 10 April 1981.

1. Paterson, M. C. & Smith, P. J. A. *Rev. Genet.* **13**, 291–318 (1979).
2. Houldsworth, J. & Lavin, M. F. *Nucleic Acids Res.* **8**, 3709–3720 (1980).
3. Painter, R. B. & Young, B. R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7315–7317 (1980).
4. Watanabe, M., Takabe, Y., Katsumata, T., Terasima, T. & Umezawa, H. *J. Antibiot., Tokyo* **26**, 417–423 (1973).
5. Makino, F. & Okada, S. *Radiat. Res.* **62**, 37–51 (1975).
6. Walters, R. A. & Hildebrand, C. E. *Biochem. biophys. Res. Commun.* **65**, 265–271 (1975).
7. Painter, R. B. & Young, B. R. *Biochim. biophys. Acta* **418**, 146–153 (1976).
8. Watanabe, I. *Radiat. Res.* **58**, 541–556 (1974).
9. Erickson, L. C., Ross, W. E. & Kohn, K. W. *Chromosoma* **74**, 125–139 (1979).
10. Dahle, D., Griffiths, T. D. & Carpenter, J. G. *Radiat. Res.* **78**, 542–549 (1979).
11. Taylor, A. M. R., Rosney, C. M. & Campbell, J. B. *Cancer Res.* **39**, 1046–1050 (1979).
12. Povrik, L. F. & Painter, R. B. *Biochim. biophys. Acta* **432**, 267–272 (1976).
13. Scudiero, D. A. *Cancer Res.* **40**, 984–990 (1980).
14. Hoar, D. I. & Sargent, P. *Nature* **261**, 590–592 (1976).
15. Suzuki, H., Nagai, K., Yamaki, H., Tanaka, N. & Umezawa, H. *J. Antibiot., Tokyo* **22**, 446–448 (1969).
16. Zambetti-Bosseler, F. & Scott, D. *Int. J. Radiat. Biol.* (in the press).
17. Taylor, A. M. R., Metcalfe, J. A., Oxford, J. M. & Harnden, D. G. *Nature* **260**, 441–443 (1976).
18. Saha, B. K. & Tolmach, L. J. *Radiat. Res.* **66**, 76–89 (1976).

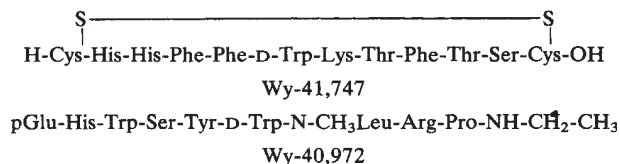
Absence of luteinizing hormone-releasing and anti-fertility properties in a glucagon-selective somatostatin analogue

Frederick J. Bex*, Alan Corbin*,
Dimitrios Sarantakis† & Eric Lien‡

Sections of Endocrinology*, Medicinal Chemistry†
and Biochemistry‡, Wyeth Laboratories, Inc., Box 8299,
Philadelphia, Pennsylvania 19101, USA

A dodecapeptide analogue of somatostatin (SRIF), [des-Ala¹, Gly²]His^{4,5}-D-Trp⁸-SRIF (Wy-41,747), combining selective inhibition of growth hormone and glucagon release with prolonged activity, prompted its consideration as an adjunct in the treatment of diabetes mellitus^{1,2}. Chronic treatment of mature rats with Wy-41,747 produced reproductive effects resembling those described for agonistic analogues of the luteinizing hormone releasing hormone (LHRH). The integral properties of the LHRH agonists such as LH release, induction of ovulation, and the ability to inhibit pregnancy both before and after implantation were noted³. We report here that the above pharmacological findings were observed only with a single batch (Wy-41,747E) which was accidentally contaminated with 0.023–0.06% by weight of an LHRH analogue (Wy-40,972).⁴

A rigorous examination of Wy-41,747E and of a second batch of Wy-41,747, designated H, was initiated after a personal communication to A.C. by J. Rivier and W. Vale concerning the absence of *in vitro* LH release with a sample of Wy-41,747 prepared and tested at the Salk Institute. The E and H batches were bioequivalent with respect to insulin and glucagon suppression in a standard procedure⁵ using male rats pretreated with Nembutal and arginine. The batches, however, were clearly differentiated by their ovulation induction activity, itself a measure of LH-releasing activity, in the Nembutal-treated proestrous rat³. The E batch induced ovulation at subcutaneous doses of 25–100 µg per rat, and such effects were dose-related, whereas Wy-41,747H had no activity at a dose of 200 µg per rat. Wy-40,972 was suspected as the most likely contaminant due to an earlier involvement in a large-scale synthesis of this potent LHRH agonist. The structures of Wy-41,747 and Wy-40,972 are:



Analytical HPLC indicated that the *k'* (retention volume) of a 0.023% impurity in Wy-41,747E was identical to the *k'* of Wy-40,972. The impurity was not detected in Wy-41,747H. Comparative binding studies in a radioimmunoassay specific for Wy-40,972 confirmed the identity of the contaminant. Wy-41,747H does not bind to the Wy-40,972 antibody, but the Wy-41,747H samples spiked with Wy-40,972 do bind to a degree related to the amounts of Wy-40,972 present and do induce ovulation in the expected manner. Binding for Wy-41,747E was evident and correlated with a 0.06% weight content of Wy-40,972.

The incident resulting in contamination has not been identified. This is the second reported case of accidental contamination with minute amounts of peptide LH-releasing agents⁶. Other investigators in this field should be cautioned to take extreme care with equipment and media previously exposed to these powerful hormones.

We conclude that the reproductive effects seen in toxicological and endocrine studies with Wy-41,747E are attributable to contamination with Wy-40,972. Wy-41,747 remains of interest as a potentially useful somatostatin analogue.

We thank Drs Jean Rivier and W. Vale of the Salk Institute, California, for drawing our attention to a possible contaminant.

Received 12 March; accepted 30 March 1981.

1. Lien, E. L., Greenwood, J. & Sarantakis, D. *Diabetes* **28**, 491–495 (1979).
2. Sarantakis, D., Teichman, J., Fenichel, R. & Lien, E. L. *FEBS Lett.* **92**, 153–155 (1978).
3. Bex, F. J., Corbin, A., Sarantakis, D. & Lien, E. L. *Nature* **284**, 342–343 (1980).
4. Foell, T. J. & Rees, R. W. U.S. Patent 4,089,946 (1978).
5. Grant, N. *et al. Life Sci.* **19**, 629–632 (1976).
6. Rivier, J. *et al. J. med. Chem.* **16**, 545–549 (1973).

Multiple sequences related to classical histocompatibility antigens in the mouse genome

B. Cami, F. Brégère, J. P. Abastado & P. Kourilsky

Unité de Biologie Moléculaire du Gène, ER CNRS 201 and SCN INSERM 20, Institut Pasteur, 28 rue du Dr Roux, 75724 Paris Cédex 15, France

A series of genes on the 17th chromosome of the mouse, called the major histocompatibility complex, has critical roles in the immune response. The classical transplantation antigens (or H-2 antigens) are encoded by genes located in several loci of this complex—*H-2K*, *H-2D* and *H-2L* (refs 1, 2; Fig. 1). The antigens are cell-surface glycoproteins consisting of a single polypeptide chain of about 350 amino acids, non-covalently

associated to a shorter polypeptide chain, β_2 -microglobulin, which is not encoded by the 17th chromosome. Within a given *H-2* haplotype, *H-2K*, *-D* or *-L* antigens differ by discrete antigenic sites, probably due to changes in their amino acid sequences³. As many as 50 alleles have been found for each *K* and *D* locus in the laboratory mouse; thus *H-2* loci are very polymorphic². We report here a preliminary study of the structure of *H-2* genes, in which the number of *H-2*-related sequences in the mouse genome has been estimated from the number of DNA fragments able to hybridize specifically with a cDNA carrying an *H-2* sequence.

We have previously described the characterization of recombinant plasmids carrying *H-2*-specific cDNA inserts^{4,5}. One of them, pH-2^d-3, was used as a source of hybridization probes to estimate the number of *H-2*, or *H-2*-related genes present in the mouse genome. A restriction map of the insert carried by pH-2^d-3 is shown in Fig. 2, together with the corresponding structural regions of the *H-2* heavy chain. pH-2^d-3 contains sequences coding for the third domain, the membrane-spanning region, the cytoplasmic COOH terminus and the 3'-untranslated region down to the poly(A) (Fig. 2).

In a first series of experiments, we used the entire pH-2^d-3 cDNA as a probe for investigating mouse genome. Liver genomic DNA was prepared from BALB/c (*H-2^d*) animals, digested by *EcoRI* and used in Southern's blotting technique⁶. Smears were observed, which were not caused by poly(A) or by G-C tails, indicating the detection by the probe of sequences highly repeated in the genome. A mouse genomic library was also screened with the same probe. This library was prepared from 15–20-kilobase BALB/c mouse DNA, produced by partial digestion with *Sau3A* and cloned in λ Charon 28 (F. Blattner, personal communication). The rate of positive response was 10–20%.

Screening experiments were then repeated using, in parallel, *PstI* restriction fragments A, B and C as probes. The 160-base pair fragment C was the only one to hybridize with 30 genomic clones out of 200, whereas the two other probes revealed no positive responders. We conclude, therefore, that fragment C contains a sequence which is repeated about 40,000-fold in the haploid genome of the mouse. Fragment C corresponds to a part of the 3'-untranslated region of the messenger and has been shown to be highly conserved in most of the cDNAs so far sequenced specifying different *H-2* alloantigens (in preparation). It thus appears to be an intriguing and original feature of the *H-2* genes that they contain a highly repeated sequence in the 525-base pair, 3'-untranslated region of their mRNA.

To probe specifically for *H-2* coding sequences, we prepared a *BglI*–*HinfI* restriction fragment (probe D, Fig. 2). It codes for amino acids 195–283 of *H-2* antigen, and corresponds nearly exactly to the third domain of this molecule⁵. Genomic DNAs of various origins were digested by the appropriate restriction enzymes and analysed using Southern's blotting technique⁶. Results are given in Fig. 3.

On hybridization with fragment D, liver BALB/c DNA (*H-2^d*), digested by *EcoRI* or *BamHI*, results in about 20 and 15 bands respectively (lanes *a* and *b*). These bands display, reproducibly, different intensities and are distributed between 1.2 and 26 kilobases. This is not peculiar to either liver DNA or BALB/c mice, because liver DNA from C57BL/6 mice (*H-2^b*), digested by *EcoRI* or *BamHI*, produced a similar number of bands (lanes *c* and *d*), as did C57BL/6 kidney DNA digested by *EcoRI* (lane *i*). The liver and kidney DNAs prepared from C57BL/6 mice yielded identical patterns (lanes *h* and *j*); in contrast, liver DNAs prepared from mice differing in their *H-2* haplotypes, resulted in different patterns (lanes *a*–*d*). Internal controls (Fig. 3, lanes *e* and *f*) rule out the possibility that these multiple bands are due to partial digestion of mouse DNA. No *EcoRI* or *BamHI* sites have been found in the cDNA sequences coding for the third domain of three different *H-2* alloantigens (ref. 5 and unpublished data). In addition, the 290-base pair probe A corresponds to the third domain of the *H-2* molecule (Fig. 2). As compared with other genes^{7,8}, introns would be

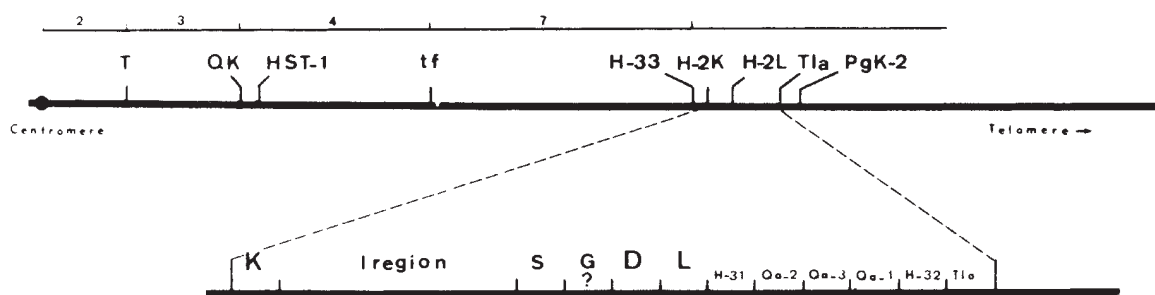


Fig. 1 Top: genetic map of the left-hand part of the 17th chromosome of the mouse, according to ref. 1. The genetic markers indicated are, from left to right: brachyury (T), quaking (QK), hybrid sterility-1 (HST-1), tufted (tf), histocompatibility 33 (H-33), histocompatibility-2K (H-2K), histocompatibility-2L (H-2L), thymusleukaemia antigen (Tla) and phosphoglycerate kinase-2 (PgK-2). Bottom: detailed genetic map of the major histocompatibility complex. Regions K, D and L, which are known to code for classical histocompatibility antigens, are indicated with thick lines. Region I contains the Ir and Ia genes, S codes for protein components of the complement, G has been reported to code for a specific H-2 antigen but its existence is controversial¹. H-31, Qa-1, -2 and -3, H-32 and Tla are minor histocompatibility antigens, sometimes referred to as 'H-2 like'. The order of the loci coding for them is not known.

Fig. 2 The restriction map of the cDNA insert of the recombinant plasmid pH-2^d-3 (ref. 5) is presented in correspondence with the linear structure of an H-2 antigen molecule¹⁶. M indicates the membrane-spanning region with the external part of the molecule on the NH₂-terminal side, and the cytoplasmic fragment on the COOH-terminal side. A, B, C and D represent restriction fragments used as radio-labelled probes in the experiments described in the text.

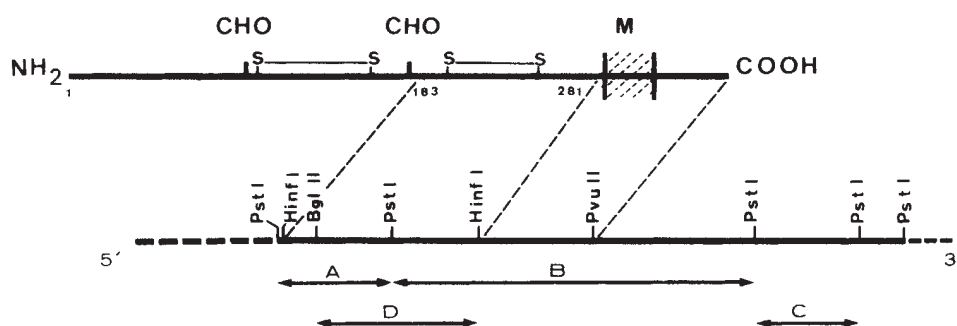


Fig. 3 The high-molecular-weight mouse genomic DNA was extracted and purified as described elsewhere¹⁷. DNA samples (20 µg) were digested by the appropriate restriction enzymes and subjected to electrophoresis in a 0.7% agarose gel. The DNA was then transferred to nitrocellulose sheets as described by Southern⁶. The sheets were prehybridized for 4 h at 68 °C in 10× Denhardt (1× Denhardt = 0.02% Ficoll, 0.02% polyvinylpyrrolidone, 0.02% bovine serum albumin), 6× SSC (1× SSC = 0.15 M NaCl, 0.015 M Na citrate, pH 7). Restriction fragment D from pH-2^d-3 (Fig. 2) was polymerized using T4 DNA ligase (Boehringer) and then ³²P-radiolabelled by nick-translation¹⁸ to a specific activity of 2 × 10⁸ c.p.m. per µg. Each nitrocellulose sheet was hybridized with 100 ng of this probe for 16 h at 68 °C in 1× Denhardt, 2× SSC, 25 mM KH₂PO₄, pH 7.2 (NaOH), 2 mM EDTA, 0.5% SDS, 50 µg ml⁻¹ sonicated herring sperm DNA. The nitrocellulose sheets were then washed four times for 45 min at 68 °C in 1× Denhardt, 2× SSC, 0.5% SDS, and twice for 45 min at 68 °C in 0.2× SSC, 0.1% SDS. They were dried for 10 min at 80 °C and autoradiographed on flash-activated Kodak Royal-XO-Mat films for 3–10 days at –70 °C with intensifying screens. Lanes a and b, liver DNA from BALB/c mice (H-2^d haplotype) digested by *Eco*RI and *Bam*HI, respectively. Lanes c and d, liver DNA from C57BL/6 mice (H-2^b haplotype) digested by *Eco*RI and *Bam*HI respectively. Lanes e and f, same as lanes 1 and 2, but 0.5 ng of pH-2^d-3 DNA has been added to mouse DNA before digestion, to provide total digestion controls. Lanes g, h and i, liver DNA from BALB/c, liver and kidney DNAs from C57BL/6 mice, respectively, were digested with *Eco*RI.

expected to occur between, rather than inside, the domains. The presence of *EcoRI* or *BamHI* sites is then unlikely to occur within the genomic sequences coding for the third domain of H-2 antigens: as a consequence, most of these genomic sequences are expected to give a single hybridization band. Hybridizations were carried out in stringent conditions (Fig. 3 legend): no cross-hybridization of the probe is thus expected to take place with sequences coding for immunoglobulin constant domains or for β_2 -microglobulin, although these display some amino acid sequence homologies with the third domain of H-2 antigens^{9,10}. Similarly, no cross-hybridization should be detected with sequences coding for the first and second domains of H-2 antigens, which share no apparent homology with the third domain^{11,12}.

The same mouse genomic library as above was screened, using as a probe fragment B, which partially extends into the third domain (Fig. 2). As cuts in mouse DNA have been generated at random and the size of the inserts is much larger than that of the probe, the frequency of positive clones should provide a measure of the number of H-2-related sequences in the mouse genome: about 750,000 recombinants were screened and 65 positive ones isolated. The library was not composed of independent clones and there could have been selective amplification of some of them. However, restriction mapping of the DNA of isolated positive clones showed that more than half of them, at least, were different from one another (data not shown). Because the mouse genome should be represented in about 250,000 recombinants, the results indicate the existence of at least 10–12 sequences reacting with probe B per haploid genome. This figure is most probably an underestimate because many positive recombinants usually escape or are lost during isolation.

Thus, the present data suggest strongly that 15–20 restriction fragments of the mouse genome possess sequences sharing extensive homology with the third domain of H-2 antigens, three to five times more than expected from the genetics of H-2 antigens. Of course, some of these sequences could correspond to pseudogenes, as described in the globin gene family^{13,14}. Also, 15–20 H-2-related sequences can exist in the mouse genome. A possible explanation for the existence of bands of stronger intensity (Fig. 3) could be the occurrence of several H-2 (third domain) genomic sequences in the corresponding DNA fragments. On the other hand, note that several cell-surface antigens, such as Tla, Qa1, -2 and other minor transplantation antigens, are composed of a heavy chain of molecular weight 44,000 associated with β_2 -microglobulin, a structure like that of H-2 antigens¹⁵. The sequence homologies reported between the third domain of H-2 molecules, β_2 -microglobulin and the constant domains of immunoglobulins suggest that structures similar to the third domain may be conserved in a variety of genes. In particular, this might be expected for genes coding for proteins interacting with β_2 -microglobulin itself.

In this respect, the use of a hybridization probe specific for the third domain may have revealed the existence of a multigene family containing, in addition to genuine H-2K, -D and -L genes, perhaps, Tl, Qa1 and/or unpreviously characterized genes encoding H-2-like antigens. Analysis of the isolated genes should answer this question.

We thank Dr F. Blattner for the gift of the mouse genomic DNA library, Dr F. Rougeon for mouse kidney DNA, Miss O. Le Bail for technical assistance, and Dr G. Gachelin for critical reading of this manuscript. The work reported was done in accordance with the French Guidelines for Recombinant Research. It was supported by grants from the CNRS (ER 201, ATP 4246), INSERM (SCN 20, CRAT 72-79-104) and DGRST (79-7-0500, 78-7-2931).

Received 11 March; accepted 8 April 1981.

- Klein, J. *Science* **203**, 516–521 (1979).
- Klein, J. *Biology of the Mouse Histocompatibility-2 Complex* (Springer, New York, 1979).
- Nairn, R., Yamaga, K. & Nathenson, S. G. *A. Rev. Genet.* **14**, 241–277 (1980).
- Kvist, S. *et al. Proc. natn. Acad. Sci. U.S.A.* **78** (in the press).
- Brégère, F. *et al. Nature* (in the press).

- Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
- Sakano, H. *et al. Nature* **277**, 627–633 (1979).
- Eaton, W. A. *Nature* **284**, 183–185 (1980).
- Tragardh, L., Rask, L., Wiman, K., Fohlman, J. & Peterson, P. A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5839–5842 (1979).
- Orr, H. T., Lopez de Castro, J. A., Lancet, D. & Strominger, J. L. *Nature* **282**, 266–270 (1979).
- Orr, H. T., Lopez de Castro, J. A., Lancet, D. & Strominger, J. L. *Biochemistry* **18**, 5711–5720 (1979).
- Tragardh, L., Rask, L., Wiman, K., Fohlman, J. & Peterson, P. A. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1129–1133 (1980).
- Proudfoot, N. J. & Maniatis, T. *Cell* **21**, 537–544 (1980).
- Lacy, E. & Maniatis, T. *Cell* **21**, 545–553 (1980).
- Vitetta, E. S. & Capra, J. D. *Adv. Immun.* **26**, 147–193 (1968).
- Kimball, E. S., Maloy, E. S., Martinko, J. M., Nathenson, S. G. & Coligan, J. E. *Molec. Immun.* **17**, 1283–1291 (1980).
- Gross-Bellard, M., Oudet, P. & Chambon, P. *Eur. J. Biochem.* **36**, 32–38 (1973).
- Maniatis, T., Jeffrey, A. & Kleid, D. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1184–1188 (1975).

Homologous tyrosine phosphorylation sites in transformation-specific gene products of distinct avian sarcoma viruses

James C. Neil*, Jacques Ghysdael*, Peter K. Vogt* & John E. Smart†

* Department of Microbiology, University of Southern California School of Medicine, 2011 Zonal Avenue, Los Angeles, California 90033, USA

† Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA

Three classes of avian sarcoma viruses have been defined recently (Table 1)^{1–5}. Transformation-specific proteins of all three viral classes have associated tyrosine-specific protein kinase activity^{4–12}. The protein kinase activity of Rous sarcoma viruses (RSV), representing class I, seems to be a property of the transformation-specific protein, pp60^{src}, itself¹³ and seems to be essential for cellular transformation, as the activity is labile or absent in transformation mutants^{6,7,14,15}. This connection between transformation-specific protein, enzyme activity and oncogenic cellular changes has yet to be established for the viruses of classes II and III, which code for polypeptides in which gag-derived and transformation-specific sequences are fused. We report here an analysis of the phosphorylation sites on the transformation-specific proteins encoded by RSV and the class III viruses which shows that, although there is no appreciable sequence relationship between large portions of these molecules, a tryptic peptide containing the phosphotyrosine seems to be identical in classes I and III.

In this study we made use of the fact that class II and III avian sarcoma virus transformation-specific proteins are labelled in *in vitro* phosphotransfer reactions on immunoprecipitates^{3–5,9–12}. Until conclusive experiments using purified viral proteins have been done, the possibility remains that these reactions are catalysed by cell-coded proteins associated with the immunoprecipitated material. Note that there are conflicting reports on

Table 1 Avian sarcoma viruses

Class	Virus	Transformation-specific protein
I	RSV Avian sarcoma virus B77	pp60 ^{src}
II	FSV Avian sarcoma virus PRCII	P140 P105
III	Avian sarcoma virus Y73 ESV	P90 P80

This classification is based on homology between transformation-specific genetic sequences^{1,2} and gene products^{3–5} of the sarcoma viruses. Avian acute leukaemia viruses which also induce sarcomas such as AEV and MC29 are omitted.

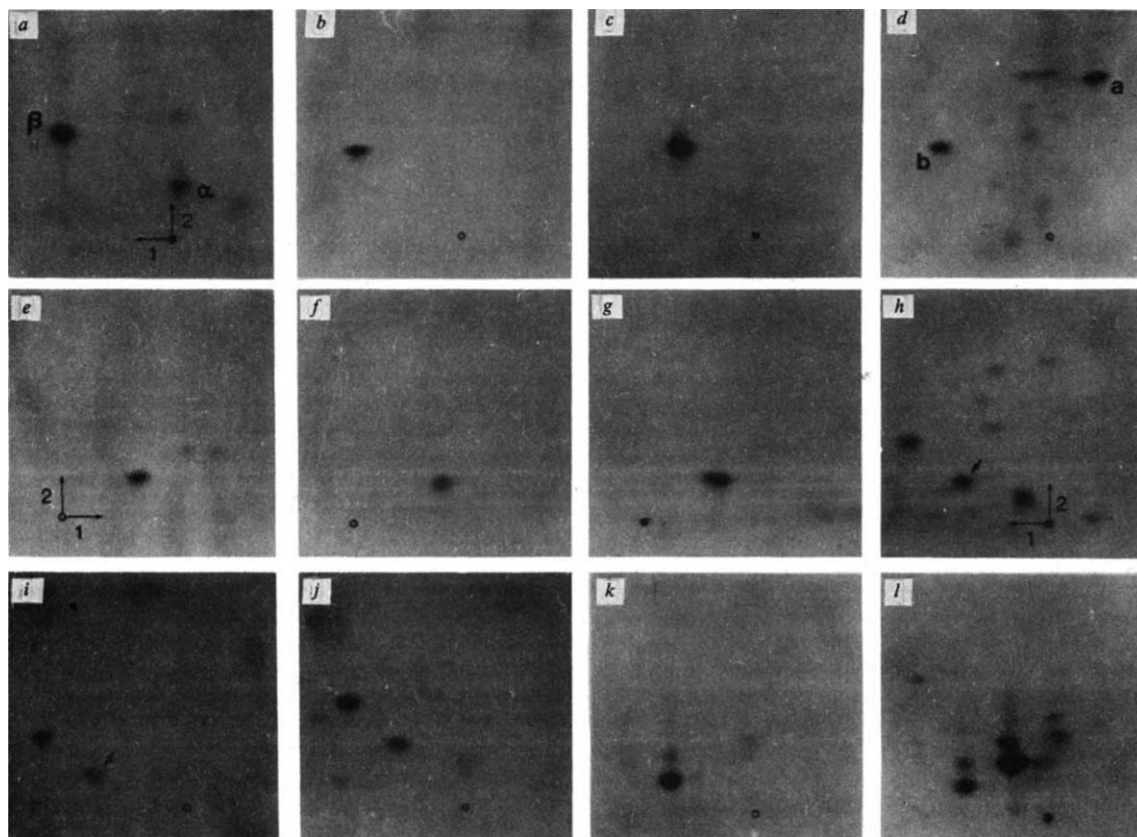


Fig. 1 Tryptic phosphopeptide analysis of RSV pp60^{src}, Y73 P90, ESV P80 and PRCII P105. Proteins were labelled either by metabolic labelling of cells with ³²P-orthophosphate in media lacking phosphate or by *in vitro* labelling in immunoprecipitates with [γ -³²P]ATP^{6,11} (isotopes from ICN). To label pp60^{src}, cells transformed by the Schmidt-Ruppin strain of RSV (SR-RSV-A) were grown overnight in the presence of ³²P-orthophosphate. In the case of ESV P80, labelling was for 2 h. As previously described²⁵, cells were lysed with detergents, and proteins immunoprecipitated with antisera to virion gag proteins (for the polyproteins) or antiserum from RSV tumour-bearing rabbits (given by Dr R. Erikson, University of Colorado). To label proteins in *in vitro* phosphotransfer reactions, the viral polyproteins were immunoprecipitated from unlabelled cells and labelled with [γ -³²P]ATP (1,000–2,000 Ci mmol⁻¹) at 30 °C for 10 min as described elsewhere¹¹. Proteins labelled metabolically ('*in vivo*') or *in vitro* were separated on SDS-polyacrylamide gels. The gels were dried and the protein bands located by autoradiography. The bands were excised, the proteins eluted, precipitated with carrier protein, washed in organic solvents, oxidized with performic acid and digested with TPCK-trypsin. This procedure has been described in detail elsewhere²⁵. The resultant peptides were separated by thin-layer electrophoresis at 1,000 V for 27 min, in a buffer containing 1% ammonium carbonate pH 8.9 (ref. 8). The second-dimension separation was chromatography in *n*-butanol/acetic acid/water/pyridine (75:15:60:60) for 4 h. Analyses in e, f and g used, instead, electrophoresis at 600 V for 60 min in a buffer containing *n*-butanol/pyridine/acetic acid/water (2:1:1:80), pH 4.5 and chromatography in *n*-butanol/pyridine/acetic acid/water (65:50:10:40) for 4 h. Phosphopeptides were detected by autoradiography on Kodak NS film at room temperature, or Kodak XR film at -70 °C with Du Pont Cronex intensifying screens. The phosphoamino acid content of phosphopeptides α and β of pp60^{src} has been reported by others⁸ and confirmed by our analyses (not shown). We have described a similar analysis of ESV P80 and Y73 P90 phosphopeptides a and b (ref. 5). Phosphopeptides a and α contain phosphoserine, whereas b and β contain phosphotyrosine. Phosphopeptide α of pp60^{src} is not seen in e or g as it migrates rapidly off the TLC plate. Phosphoamino acid analyses confirmed that the new V8 cleavage product peptides contained phosphotyrosine (not shown). The maps in a–g and k are of tryptic digests; the phosphoproteins mapped in h–j and l were digested with TPCK-trypsin followed by staphylococcal V8 protease. The maps are of a, pp60^{src} labelled *in vivo*; b, ESV P80 labelled *in vitro*; c, Y73 P90 labelled *in vitro*; d, Y73 P90 labelled *in vivo*; e, pp60^{src} labelled *in vivo*; f, Y73 P90 labelled *in vitro*; g, a mixture of e and f; h, pp60^{src} labelled *in vivo*; i, ESV P80 labelled *in vitro*; j, a mixture of h and i; k, PRCII P105 labelled *in vitro*, and l, PRCII P105 labelled *in vitro*.

the ability of partially purified pp60^{src} to phosphorylate itself^{16,17}.

In a previous analysis of the tyrosine phosphorylation sites of PRCII and Y73 (avian sarcoma viruses) and FSV (Fujinami sarcoma virus) proteins³, we noted that the grouping of these viruses, based on nucleic acid hybridization and peptide mapping, was reflected in the phosphopeptides. We have also shown that the phosphorylation sites on the class III virus proteins are indistinguishable, as might be expected for such closely related viruses⁵. Analysis of pp60^{src} of RSV showed that the tyrosine-containing tryptic peptide of pp60^{src} co-migrated with the equivalent peptides of the Y73 and ESV (Esh sarcoma virus) polyproteins (Fig. 1a–c). Like pp60^{src}, Y73 P90 and ESV P80 have additional serine phosphorylation sites that are not found in the helper virus gag gene precursor protein Pr76 (Fig. 1d). However, the serine phosphorylation sites are not labelled in *in vitro* phosphotransfer reactions (Fig. 1)⁵. Our primary interest is in the *in vitro* kinase target sites, as these are likely to be related to oncogenic function^{6,7,14,15}. We analysed the tyrosine-containing tryptic peptides further to discover whether the co-migrating peptides were really identical or showed fortuitous similarity only in the separation conditions of the initial

experiment. As a first step we compared electrophoretic migration at a different pH, and using a different chromatography buffer (Fig. 1e–g). Co-migration was still observed in these conditions, which suggests that the amino acid composition of these phosphotyrosine-containing peptides is very similar.

Separation of the phosphotyrosine-containing tryptic peptides of Y73 P90 and ESV P80 by HPLC on a C18 reverse-phase column shows that the major ³²P-labelled peptide of each protein elutes at the same point in the acidic ethanol gradient; the major phosphotyrosine-containing tryptic peptide of pp60^{src} elutes at the same position in these conditions (data not shown, see (ref. 18)). Sequential Edman degradation of these ³²P-labelled peptides shows that the phosphotyrosine occurs at residue 7 in both Y73 and ESV peptides (Fig. 2), as well as in pp60^{src} (ref. 18).

The DNA sequence of the RSV *src* gene has been determined, and these data have been used to deduce a possible amino acid sequence for pp60^{src} (ref. 19). Partial sequence analysis of the methionine-labelled tryptic peptides of pp60^{src} has demonstrated that the predicted amino acid sequence is correct over most of the length of the protein¹⁸. Using similar techniques, the amino acid sequence of the phosphotyrosine-

containing tryptic peptide of pp60^{src} was deduced: N-Leu-Ile-Glu-Asp-Asn-Glu-Tyr-Ala-Arg-C. The predicted amino acid sequence of this peptide in pp60^{src} includes potential cleavage sites for staphylococcal V8 protease, which cleaves at the carboxy-terminal side of glutamic acid residues²⁰. We therefore digested pp60^{src} and ESV p80 with TPCK-trypsin, followed by V8 protease. In each case there was partial cleavage of the phosphotyrosine-containing tryptic peptide, generating a new peptide. Co-migration of the newly generated peptides (Fig. 1*h-j*) suggested that the cleavage products are identical in pp60^{src} and ESV P80. Similar results were obtained with P90 of Y73. Complete V8 protease digestion of the phosphotyrosine-containing tryptic peptide of ESV P80 yields a single new peptide in which the phosphotyrosine occurs at residue 1 (data not shown). Thus we suggest that a tryptic decapeptide containing the major tyrosine phosphorylation site of pp60^{src} is identical in the class III avian sarcoma virus polyproteins.

As the RSV *src* gene and the transformation-specific sequences of Y73 virus have no homology detectable by nucleic acid hybridization², and because the polyproteins of Y73 and ESV show no overlap with pp60^{src} in their methionine-containing tryptic peptides³⁻⁵, it is assumed that the region of amino acid homology around the tyrosine phosphorylation sites represents a discrete and limited domain. We are now examining the extent of the sequence homology around these sites by cleavage with other proteases.

In addition to the homology in kinase target sites of pp60^{src}, P80 and P90 which we have described here, there is evidence that some features of these sites also occur in other transformation-specific proteins. We have found that the tyrosine phosphorylation sites on the class II avian sarcoma virus transformation-specific proteins, although not identical to classes I and III, are on tryptic peptides that are susceptible to digestion with V8 protease, yielding multiple partial cleavage products (Fig. 1*k,l*). Furthermore, polyoma virus middle-T antigen also has an associated protein kinase activity²¹ which phosphorylates a tyrosine residue²². It is striking that the proposed sequence of middle-T contains a sequence of six glutamic acid residues followed by tyrosine; this tyrosine has been identified as the likely site of phosphorylation in the *in vitro* reaction (A. E.

Smith, personal communication). These data and our results suggest that the occurrence of acidic amino acids adjacent to the phosphoacceptor tyrosine may be a general property of target sites for tyrosine-specific protein kinases, just as basic amino acid residues usually occur at serine phosphorylation sites²³. These observations also support the hypothesis that such conserved sites may have a central role in regulating the function(s) of the viral transforming proteins as well as their normal cellular homologues.

During the course of this work we learned of similar observations on the phosphorylation sites of the Y73 and RSV proteins²⁴.

This work was supported by USPHS grants CA 13213 and CA 19725. J. N. holds a postdoctoral fellowship from the Leukemia Society of America. J. G. is supported by postdoctoral fellowships from the Fondation Rose et Jean Hoguet and NATO.

Received 27 January; accepted 9 April 1981.

- Shibuya, M., Hanafusa, T., Hanafusa, H. & Stephenson, J. R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6536-6540 (1980).
- Yoshida, M., Kawai, S. & Toyoshima, K. *Nature* **287**, 653-654 (1980).
- Neil, J. C., De Lamarier, J. & Vogt, P. K. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1906-1910 (1981).
- Beemon, K. *Cell* **24**, 145-154 (1981).
- Ghydael, J., Neil, J. C. & Vogt, P. K. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2611-2615 (1981).
- Collett, M. S. & Erikson, R. L. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2021-2024 (1978).
- Levinson, A. D., Oppermann, H., Levintow, L., Varmus, H. E. & Bishop, J. M. *Cell* **15**, 561-572 (1978).
- Hunter, T. & Sefton, B. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1311-1315 (1980).
- Kawai, S. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 6199-6203 (1980).
- Feldman, R., Hanafusa, T. & Hanafusa, H. *Cell* **22**, 757-765 (1980).
- Neil, J. C., Ghydael, J. & Vogt, P. K. *Virology* **109**, 223-228 (1981).
- Ghydael, J., Neil, J. C., Wallbank, A. M. & Vogt, P. K. *Virology* **111**, 386-400 (1981).
- Erikson, R. L., Collett, M. S., Erikson, E. & Purchio, A. F. *Proc. natn. Acad. Sci. U.S.A.* **12**, 6200-6204 (1979).
- Sefton, B. M., Hunter, T. & Beemon, K. *J. Virol.* **33**, 220-229 (1980).
- Oppermann, H., Levinson, A. D. & Varmus, H. E. *Virology* **108**, 47-70 (1980).
- Erikson, R. L., Collett, M. S., Erikson, E., Purchio, A. F. & Brugge, J. S. *Cold Spring Harb. Symp. quant. Biol.* **44**, 907-917 (1979).
- Levinson, A. D., Oppermann, H., Varmus, H. E. & Bishop, J. M. *J. biol. Chem.* **255**, 11973-11980 (1980).
- Smart, J. E., Oppermann, H., Czernilofsky, A. P. & Bishop, J. M. *Proc. natn. Acad. Sci. U.S.A.* (submitted).
- Czernilofsky, A. P. *et al. Nature* **287**, 198-203 (1980).
- Houmard, J. & Drapeau, G. R. *Proc. natn. Acad. Sci. U.S.A.* **69**, 3506-3510 (1972).
- Smith, A. E., Smith, R. & Paucha, E. *Cell* **18**, 915-924 (1979).
- Eckhart, W., Hutchinson, M. A. & Hunter, T. *Cell* **18**, 925-933 (1979).
- Krebs, E. G. & Beavo, J. A. *Rev. Biochem.* **48**, 923-959 (1979).
- Patschinsky, T. & Sefton, B. M. (in preparation).
- Neil, J. C., Breitman, M. & Vogt, P. K. *Virology* **108**, 98-110 (1981).
- Smart, J. E. & Ito, Y. *Cell* **15**, 1427-1437 (1979).
- Smart, J. E., Lewis, J. B., Mathews, M. B., Harter, M. L. & Anderson, C. W. *Virology* (in the press).

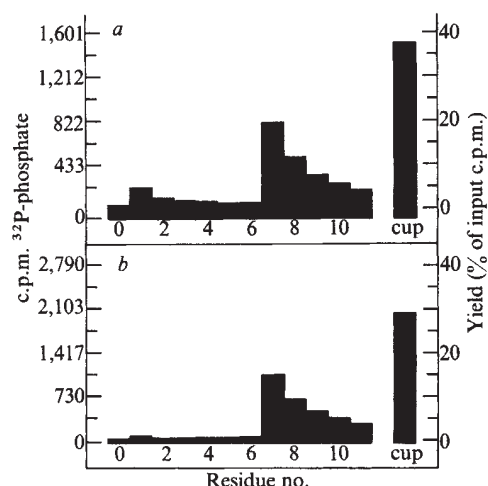


Fig. 2 Automatic sequential Edman degradation of the ³²P-labelled tryptic peptides of polyproteins of Y73 and ESV avian sarcoma viruses. Tryptic peptides were first separated on a Sepheriso-b (C18) reverse-phase HPLC column. *a*, Peptide eluting at position 50-58 from Y73 P90 and *b*, peptide eluting at position 52-62 from ESV P80. ³²P-labelled peptides in 2 mg of apomyoglobin were dissolved in 500 µl of 18% formic acid. A 50-µl aliquot was taken for counting and the remainder applied to the spinning cup of a Beckman 890C sequencer with 3 mg of polybrene (Sigma). Sequencing was performed using a 0.1 M Quadrol buffer program. Conversion of the anilinothiazolinone derivatives, counting and determination of the carrier apomyoglobin residues (to ascertain that the sequencer had functioned properly) were as described by Smart *et al.*²⁷. A zero cycle was included in each run by omitting the heptafluorobutyric acid cleavage step during the first programmed cycle. Recovery of the ³²P-labelled amino acid derivatives is incomplete on each cycle because of their strong affinity for polybrene.

The structure of a chromosomal leghaemoglobin gene from soybean

Erik Østergaard Jensen, K. Paludan, J. J. Hyldig-Nielsen, P. Jørgensen & K. A. Marcker

Department of Molecular Biology and Plant Physiology, University of Aarhus, DK-8000 Århus C, Denmark

Leghaemoglobin (Lb) has only been found in the nitrogen-fixing root nodules of legumes symbiotically associated with *Rhizobia*. Soybean nodules contain three major forms of Lb called *a*, *c*₁ and *c*₂, respectively. The amino acid sequences of the three Lbs have been determined recently¹. Lb *a* differs from the two *c* varieties in six amino acids, while Lb *c*₁ and *c*₂ differ in only one amino acid. We have recently reported the isolation of a hybrid plasmid containing soybean Lb cDNA sequences². Using one such clone as a probe, we have isolated five different soybean DNA fragments containing chromosomal Lb genes from a soybean DNA library which was constructed using lambda Agt Wes-λB as vector. The detailed analysis of one of the isolated chromosomal Lb genes revealed the presence of three intervening sequences. Alignment by homology of all known globin amino sequences shows that the splicing points of the first and the third intervening sequences in the Lb-coding sequence are located in the same positions as the two interruptions found in all other known globin-coding sequences.

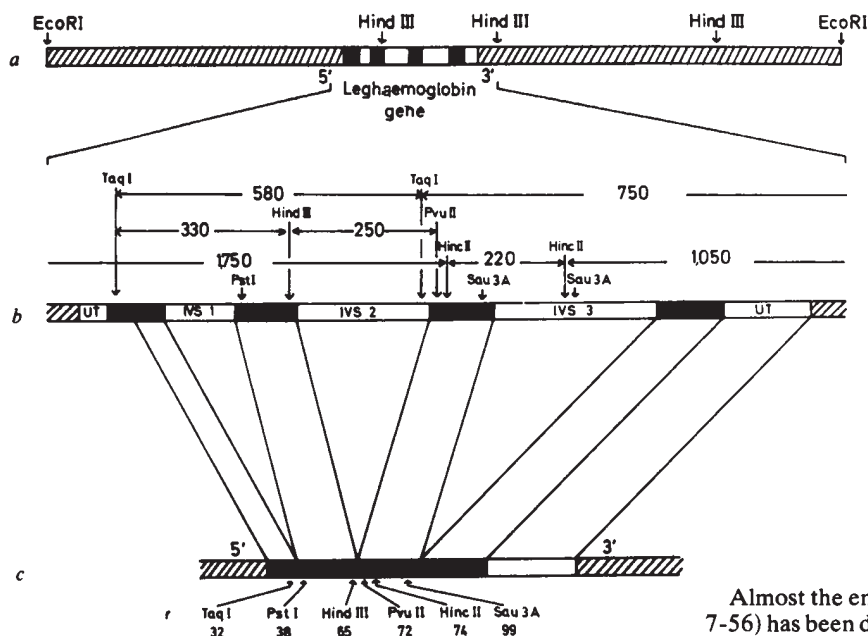


Fig. 1 Restriction cleavage map of the 7.5-kilobase soybean DNA fragment containing a Lb gene. Solid and open boxes represent coding and noncoding sequences, respectively; cross-hatched boxes represent non-Lb gene DNA sequences. UT 5' and UT 3' represent the sequence corresponding to the non-translated 5' and 3' ends of the Lb mRNA. IVS-1, IVS-2 and IVS-3 denote intervening sequences. Restriction endonuclease cleavage sites are indicated by arrows, and the distances between cleavage sites are given in base pairs. *a*, Map of the cloned *EcoRI* fragment of soybean DNA including the Lb gene. *b*, Expanded map of the Lb gene. *c*, Map of the Lb cDNA clone 7-56. Restriction endonuclease sites are here correlated with the number of the appropriate codon in the mRNA sequence. Cross-hatched boxes represent pBR322 DNA sequences.



Fig. 2 Electron micrograph of the hybrid between Lb mRNA and the chromosomal Lb gene, formed by incubating the DNA and RNA in a buffer containing 80% formamide, 0.1 M HEPES pH 7.2, 0.4 M NaCl and 0.01 M EDTA for 2 h at 51 °C (ref. 15). The mixture was then diluted 10-fold into 30% formamide, 0.1 M Tris-HCl pH 8.5, 0.1 mg ml⁻¹ cytochrome *c* and spread onto a water hypophase. After staining and shadowing with platinum, pictures were taken in a Philip EM 301. The arrows delineate the mRNA-DNA hybrid. 1, 2 and 3 refer to IVS-1, IVS-2 and IVS-3, respectively.

Almost the entire DNA sequence of a Lb cDNA clone (called 7-56) has been determined, lacking only that part corresponding to the first 15 amino acids of the Lb amino acid sequence. The DNA sequence is consistent with the amino acid sequence determined by Sievers *et al.*¹ for Lb c except that (1) Sievers *et al.*¹ determined the sequence -Phe-Val-Val-Lys- at amino acid positions 102-105 whereas we consistently find in this region a sequence corresponding to -Phe-Val-Val-Lys-; and (2) Sievers *et al.*¹ found either Phe or Lys as a -COOH terminal whereas our equivalent base sequence corresponds to Ala.

Southern blotting analysis³ of *EcoRI* digests of soybean DNA revealed the presence of at least seven different Lb genes in soybean, of length 1.4, 4.2, 5.5, 6.0, 7.5, 12 and 13 kilobases, respectively. Essentially similar results have recently been reported by Sullivan *et al.*⁴.

To investigate the chromosomal Lb genes in more detail a library of soybean genes was constructed using λ gt Wes- λ B as cloning vector⁵. Soybean DNA was digested to completion with *EcoRI* and ligated to *EcoRI*/*Sst*-digested λ gt Wes- λ B. About 8×10^5 recombinant phage were screened⁶. Several plaques giving positive signals were purified and the DNA thus obtained digested with *EcoRI* and subjected to Southern analysis using nick-translated 7-56 DNA as a probe. In this way five separate clones carrying chromosomal Lb genes were detected. The sizes of the cloned soybean DNA fragments were 1.4, 4.2, 7.5, 12 and 13 kilobases, respectively. The clone containing the 7.5-kilobase fragment was subjected to further analysis by digestion with a variety of restriction endonucleases, from which a cleavage map for this particular Lb DNA gene could be constructed, as shown in Fig. 1. The sizing of the various restriction fragments suggested that the coding Lb DNA sequence was interrupted in three places with intervening sequences. To confirm this the 7.5-kilobase fragment was subcloned into pBR322, and the resulting clone allowed to form R-loops in the presence of Lb mRNA. The mixture was then analysed in the electron microscope with the result shown in Fig. 2; three loops are clearly indicated. The sizes of the three intervening sequences were estimated by direct measurement in the electron microscope and by sizing of the appropriate fragments obtained by digestion of Lb DNA with various restriction endonucleases. The results indicate that IVS-1 is ~160 base pairs long, IVS-2 230 base pairs and IVS-3 300 base pairs.

To establish the positions of the splicing points the 7.5-kilobase fragment was digested with *Sau3A* and the resulting fragments were subcloned into the M13 mp2 *Bam* vector developed by Messing⁷. Clones containing structural Lb DNA sequences were located by plaque hybridization using nick-translated Lb cDNA from clone 7-56. M13 clones containing Lb coding sequences were then subjected to DNA sequence analysis using the dideoxy chain-termination method developed by

Sanger *et al.*⁸. As the coding Lb sequence contains a *Hind*III cleavage site in the position corresponding to amino acids 64–65, the 5' ends resulting from a *Hind*III digestion of the 7.5-kilobase fragment could be labelled with ³²P using [γ -³²P]ATP and T₄ polynucleotide kinase. Appropriate DNA fragments were then subjected to DNA sequence analysis according to Maxam and Gilbert⁹.

Figure 3 shows the sequences determined around the splicing junctions. In the Lb gene the coding sequence is interrupted at the DNA sequences corresponding to the following amino acids in the amino acid sequence; IVS-1 at amino acid 32, IVS-2 between amino acids 68 and 69 and IVS-3 between amino acids 103 and 104. The sequences around the splicing junctions are consistent with the Breathnach–Chambon¹⁰ rule because intervening sequences start with GT- and terminate with -AG. Furthermore, the sequences at the 5' and 3' ends of the intervening sequences agree with the consensus sequence for splicing junctions suggested by Lewin¹¹. Thus plant genes share this feature with other eukaryotic genes¹².

Two intervening sequences have been found in all globin genes so far examined; in α globins the first interrupts between

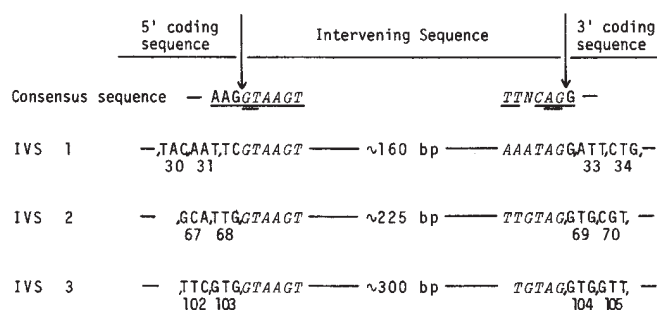


Fig. 3 Comparison of the position of Lb intervening sequence with that of a consensus sequence from Lewin¹¹. Single underlines indicate a frequency of occurrence $\geq 45\%$, double underlines a frequency $\geq 95\%$. Numbers under codons refer to the number of the respective amino acid in the protein sequence. Codon 31 signifies asparagine, which identifies the isolated gene as a Lb c gene¹. The DNA sequences were determined by the dideoxy chain-termination method described by Sanger *et al.*⁸ and the Maxam–Gilbert technique⁹.

codons 31 and 32 and the second between codons 99 and 100. For β -globins the corresponding two intervening sequences are located between codons 30 and 31, and 104 and 105, respectively. In the Lb gene, interruptions occur at codons 32 and between 103 and 104.

By computer analysis Hunt *et al.*¹³ have compared homologies of the amino acid sequences of globins including several Lbs. The observed similarity between the structures suggested that Lb and other globins have a common evolutionary origin, although they could be the result of convergent evolution. Inspection of the homology alignment of Hunt *et al.*¹³ reveals that the coding sequences of all known globin genes, including the Lb gene, are interrupted by intervening sequences at identical positions. This finding supports the notion that all globin genes originate from a common ancestral gene.

The Lb gene is also interrupted between codons 68 and 69. This particular intervening sequence has not been found in the globin genes examined so far. Gilbert¹⁴ has suggested that the presence of intervening sequences in DNA can speed evolution by allowing novel proteins to be constructed from the pieces of existing ones. Thus in the case of globin genes the central coding sequence codes for the parts of the globin chain which form the haem crevice. This particular correlation between coding sequence and function is not immediately apparent for the Lb gene, where the central coding sequence is split into two pieces of almost identical length. No known functions can be assigned to the coding sequences thus generated and consequently our result does not support the attractive hypothesis that a cor-

relation always exists between a coding sequence and a particular function.

Being functional monomers, Lb and myoglobin are generally assumed to have evolved earlier from the common ancestral gene than α - and β -globins. Thus, the central intervening sequence either was present in the ancestral globin gene and subsequently lost during its evolution into α - and β -globins, or its insertion is a relatively new development in the evolution to Lb. Comparison of the Lb gene structure with that of myoglobin would help resolve this alternative but the DNA sequence of the latter remains unknown. In addition, several other Lb genes exist and their structures might in fact be different from the one determined here. Accordingly, we have cloned four other Lb genes and are investigating their detailed structures.

This work was supported by the Danish Research Council, NOVO Industry A/S and the Olga and Esper Boel Fond. We thank Drs B. Borkhardt and H. Delius for performing R-loop analysis, and Mrs Y. Gregersen and Mrs A. Marcker for technical assistance.

Received 2 February; accepted 9 April 1981.

1. Sievers, S. G., Huhtala, M.-L. & Ellfolk, N. *Acta chem. scand.* **B32**, 380–386 (1978).
2. Truelsen, E., Gausing, K., Jochimsen, B., Jørgensen, P. & Marcker, K. A. *Nucleic Acids Res.* **6**, 3061–3072 (1979).
3. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
4. Sullivan, D., Brisson, N., Goodchild, B., Verma, D. P. S. & Thomas, D. Y. *Nature* **289**, 516–518 (1981).
5. Leder, P., Tiemeier, D. & Enquist, L. *Science* **196**, 175–177 (1977).
6. Maniatis, T. *et al.* *Cell* **15**, 687–701 (1978).
7. Gronenborn, B. & Messing, J. *Nature* **272**, 375–377 (1978).
8. Sanger, F., Coulson, A. R., Barell, B. G., Smith, A. J. H. & Roe, B. A. *J. molec. Biol.* **143**, 161–178 (1980).
9. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560–564 (1977).
10. Breathnach, R., Benoist, C., O'Hara, K., Gannon, F. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4853–4857 (1978).
11. Lewin, B. *Cell* **22**, 324–326 (1980).
12. Sun, S. M., Slightom, J. L. & Hall, T. C. *Nature* **289**, 37–41 (1981).
13. Hunt, T. L., Hurst-Calderone, S. & Dayhoff, M. O. *Atlas of Protein Sequence and Structure* **5**, Suppl. 3, 229–251 (1978).
14. Gilbert, W. *Nature* **271**, 501 (1978).
15. Chow, L. T., Roberts, J. M., Lewis, J. B. & Broker, T. R. *Cell* **11**, 819–836 (1977).

A human proinsulin variant at arginine 65

D. C. Robbins*, P. M. Blix*, A. H. Rubenstein*, Y. Kanazawa†, K. Kosaka† & H. S. Tager*‡

* Departments of Biochemistry and Medicine, The University of Chicago, Chicago, Illinois 60637, USA

† Third Department of Medicine, The University of Tokyo, Japan

Human diabetes can have several causes among which are the biosynthesis and secretion of mutant insulins with defective function. We recently described a patient whose diabetes resulted from secretion of an insulin in which leucine substituted for phenylalanine at position B24 (refs 1–4). Studies showing allelic variation in untranslated and intervening sequences of the human insulin gene (even in normal individuals) have documented a further and unexpected degree of variation in insulin gene structure^{5,6}. We now describe the molecular abnormality of hyperproinsulinaemia in which large amounts of a proinsulin-like peptide are secreted^{7–10}. The circulating proinsulin-like material from patients in one of the two known families with hyperproinsulinaemia^{9,10} turns out to be a biosynthetic intermediate of proinsulin conversion in which the C peptide remains joined to the insulin A chain (see refs 11–13 for reviews of proinsulin biosynthesis and conversion). An amino acid alteration at position 65 (a position normally filled by arginine) apparently blocks conversion of the intermediate to biologically active insulin.

§ Present address: Department of Medicine, University of Vermont, Burlington, Vermont 05401, USA.

‡ To whom correspondence should be addressed.

Sera from the three affected members of the family (a 62-yr-old man with glucose intolerance, his 57-yr-old sister and his 42-yr-old son) all showed marked elevations in fasting concentrations of immunoreactive insulin ($70\text{--}100\ \mu\text{U ml}^{-1}$ compared with normal values of $<25\ \mu\text{U ml}^{-1}$) and abnormally high ratios of proinsulin-like material to insulin ($9\text{--}10$ compared with normal values of <0.25 , ref 14). Insulin-related components from the sera of the two younger patients were purified by immunoaffinity chromatography^{2,15} using guinea pig anti-porcine insulin serum pooled from several immunized animals. Gel filtration of the eluted peptides on Bio-Gel P-30 (equilibrated with 3 M acetic acid containing $5\ \mu\text{g ml}^{-1}$ of bovine serum albumin) resulted in the separation of the 9,000-molecular weight (MW) proinsulin-like peptide from the smaller amount of 6,000-MW insulin. In all, 400 ml of serum was

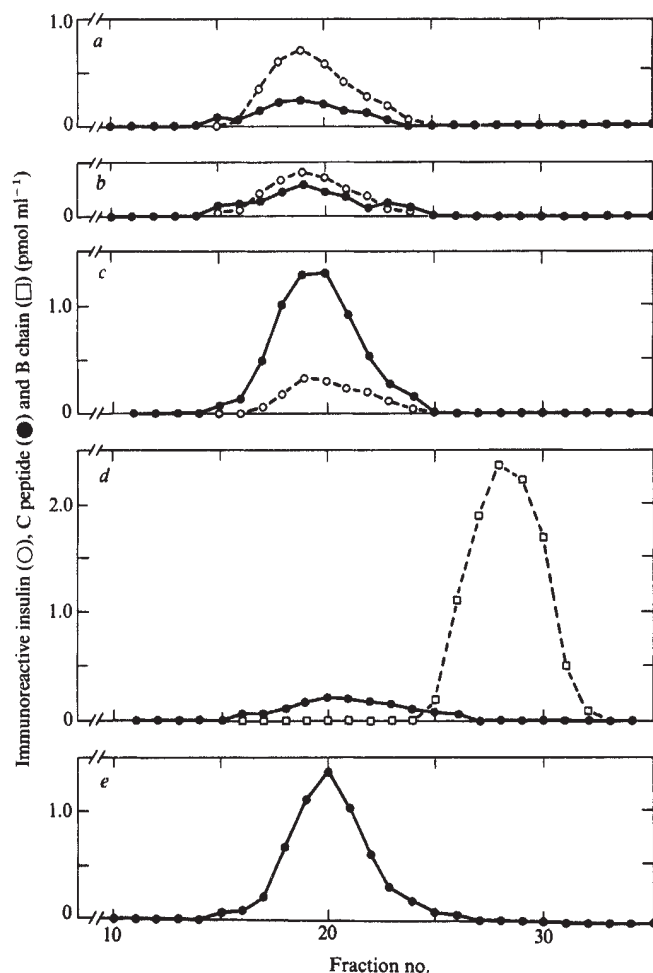


Fig. 1 Gel filtration and immunoreactivity of purified proinsulin-like material before and after oxidative sulphytolysis. The column of Bio-Gel P-10 ($0.9 \times 90\text{ cm}$) was equilibrated with a buffer containing 0.1 M Tris brought to pH 7.6 with HCl, 0.05 M NaCl, 0.2 mg ml^{-1} sodium azide and $5\ \mu\text{g ml}^{-1}$ bovine serum albumin; 1.2-ml fractions were collected. Aliquots of column fractions were assayed for immunoreactive C peptide using antiserum M1230¹⁸ (●), insulin using antiserum GP-I¹⁶ (○) and B chain using an antiserum donated by P. R. Varadani¹⁷ (□). Human C peptide, porcine insulin and porcine bis-S-sulpho B chain served as the respective immunoassay standards. *a*, Native peptide, untreated fractions; *b*, native peptide, trypsin-digested fractions [$100\ \mu\text{g ml}^{-1}$ tosylphenylalanine chloromethyl ketone-treated trypsin (Worthington), 10 mM CaCl_2 , 60 min, 37°C , briefly heated to 100°C after digestion]; *c*, native peptide, fractions sequentially digested by trypsin and carboxypeptidase B [trypsin digestion as above, $100\ \mu\text{g ml}^{-1}$ carboxypeptidase B (Boehringer-Mannheim), 60 min, 37°C , briefly heated to 100°C after digestion]; *d*, oxidatively sulphytolysed peptide [1 ml of 0.2 M sodium phosphate buffer at pH 7.4 containing 8 M urea, 1.75% (w/v) sodium tetrathionate and 3.5% (w/v) sodium bisulphite, 3 h, 37°C], untreated column fractions; *e*, oxidatively sulphytolysed peptide, fractions sequentially digested by trypsin and carboxypeptidase B. Approximately 7 pmol of peptide was gel-filtered in each case. Native proinsulin and authentic bis-S-sulpho B chain appeared in fractions 19 and 28, respectively, during standardization of the column.

processed in this manner, yielding ~ 150 pmol of the proinsulin-like component; recoveries averaged 90%.

Because of the small amounts of proinsulin-like material obtainable from serum, our approach to examining the structure of the 9,000-MW peptide relied on sensitive radioimmunoassays for insulin¹⁶, insulin B chain¹⁷ and proinsulin C peptide¹⁸ to detect picomole quantities of material and on specific enzymatic and chemical manipulations to degrade the peptide to known components. Figure 1*a* shows that the purified and gel-filtered 9,000-MW peptide reacts, on a molar basis, only one-third as well with C-peptide antibodies as it does with insulin antibodies. Treatment of the column fractions with trypsin slightly enhanced the C-peptide immunoreactivity and diminished the insulin immunoreactivity (Fig. 1*b*). Sequential treatment with trypsin and carboxypeptidase B, however, resulted in a fivefold enhancement of C-peptide immunoreactivity (Fig. 1*c*). The known sensitivity of our C-peptide antiserum to the COOH-terminal sequence of human C peptide¹⁹ and the specific effect of carboxypeptidase B suggest a masking of the COOH-terminal region of the C peptide within the 9,000-MW form.

To determine whether the 9,000-MW peptide was a two-chained intermediate of conversion or single-chained proinsulin, the peptide was subjected to oxidative sulphytolysis, a procedure which irreversibly breaks both interchain and intrachain disulphide bonds²⁰. As shown in the gel filtration profile of Fig. 1*d*, the oxidized peptide separated into two distinct components, a higher-molecular-weight peptide exhibiting C-peptide immunoreactivity and a lower-molecular-weight peptide co-eluting with native B chain and showing B-chain immunoreactivity. Although treatment of individual column fractions with trypsin slightly augmented the C-peptide immunoreactivity and obliterated B-chain immunoreactivity (not shown), sequential treatment with trypsin and carboxypeptidase B, as before, enhanced the higher-molecular-weight C-peptide immunoreactivity fivefold (Fig. 1*e*). The experiments illustrated in Fig. 1 identify the patients' proinsulin-like peptide as a two-chained intermediate of conversion in which the C peptide remains joined to the insulin A chain (des-Arg 31, Arg 32-proinsulin).

We examined the proinsulin intermediate and the products of its enzymatic digestion in greater detail by polyacrylamide gel electrophoresis at pH 8.7 (ref. 21). As shown in Fig. 2*a*, the anodal mobility of the native peptide was greater than that of human proinsulin, a result consistent with the excision of basic amino acid residues from proinsulin during formation of the conversion intermediate. Digestion of the peptide with trypsin resulted in a more rapidly migrating fragment with C-peptide immunoreactivity (Fig. 2*b*). However, only sequential treatment of the intermediate with trypsin and carboxypeptidase B resulted in a fragment co-migrating with native human C peptide and showing a major enhancement in C-peptide immunoreactivity (Fig. 2*c*). As indicated by Fig. 2*d*, reversing the order of digestion with trypsin and carboxypeptidase B failed to yield native human C peptide. Analysis of the gels of Fig. 2*a-d* for insulin-related peptides showed that the insulin and C-peptide immunoreactivities of the undigested 9,000-MW peptide migrated to the same position, but that treatment with trypsin alone and with trypsin plus carboxypeptidase B yielded peptides co-migrating with desoctapeptide insulin (lacking residues B23–B30) and desnonapeptide insulin (lacking residues B22–B30), respectively (data not shown).

Further analysis of the proinsulin intermediate for an amino acid substitution in its C peptide–A chain connecting region proceeded from selective protection of amino groups by acetylation. We reasoned that acetylation would not affect tryptic cleavage of the C peptide from an intermediate lacking Lys 64, whereas acetylation would prevent removal of the C peptide from an intermediate lacking Arg 65. As shown in Fig. 2*e* (solid line), polyacrylamide gel electrophoresis of the trypsin- and carboxypeptidase B-digested, acetylated intermediate yielded a product migrating further than the native peptide (as expected

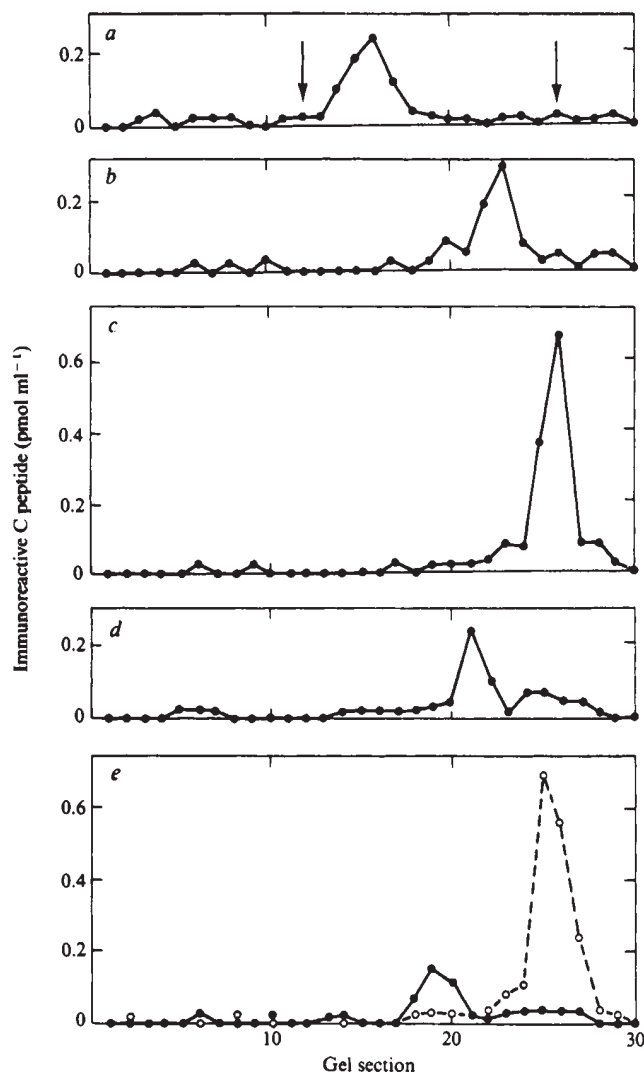


Fig. 2 Polyacrylamide gel electrophoresis of purified proinsulin-like material before and after enzymatic digestion and acetylation. Enzymatic digestions proceeded as described for Fig. 1 except that the buffer was 0.1 M *N*-ethylmorpholine adjusted to pH 8 with acetic acid; after digestion the buffer and solvent were removed under vacuum and the residue was dissolved in 0.1 ml of electrophoresis buffer before electrophoresis on 15% (w/v) acrylamide gels at pH 8.7 (ref. 21). Peptides from 1.5-mm sections of the gel were eluted into 0.5 ml of buffer¹⁸ for 20 h at 4 °C and aliquots of the resulting solutions were subjected to the C-peptide radioimmunoassay¹⁸. *a*, Native peptide; *b*, peptide digested by trypsin; *c*, peptide sequentially digested by trypsin and carboxypeptidase B; *d*, peptide sequentially digested by carboxypeptidase B and trypsin; *e*, peptide treated with acetic anhydride [four 5- μ l aliquots over a 2-h period at 22 °C in 1 ml of 2 M sodium acetate (solid line) or in 1 ml of 2 M ammonium acetate (dashed line)] before dialysis against 3 M acetic acid, evaporation and sequential digestion by trypsin and carboxypeptidase B. Approximately 2 pmol of peptide was used in each case. The left and right arrows in panel *a* show the positions taken by native human proinsulin and human C peptide, respectively. The tracking dye (bromophenol blue) appeared in gel section 30.

for a product having the C peptide joined to the A chain of desnonapeptide insulin), but to a position well behind that taken by authentic human C peptide. Control studies showed that (1) the intermediate treated with acetic anhydride in the presence of ammonium ion (to prevent amino group modification) and then subjected to purification and analysis yielded the native C peptide (Fig. 2*e*, dashed line), and (2) enzymatic digestion of acetylated human proinsulin yielded a peptide (ϵ -*N*-acetyl-C-peptidyllysine) having the electrophoretic mobility of native human C peptide (not shown).

The above results demonstrate both the presence of Lys 64 in the region joining the C peptide to the insulin A chain in the proinsulin intermediate and the presence of a nearby amino acid

alteration. Although mutations resulting in a substitution or deletion at Arg 65 [yielding the connecting region sequence Gln-Lys-(X)-Gly] or resulting in a proline for glycine substitution at position A1 (yielding the sequence Gln-Lys-Arg-Pro) could both account for this sensitivity, the latter possibility can be excluded: tryptic digestion of an intermediate with the latter sequence would have resulted in arginyl desoctapeptide insulin rather than the desoctapeptide insulin identified on the gel of Fig. 2*b*. We thus conclude that the biosynthetic intermediate results from failure of cellular enzymes to process correctly a proinsulin variant having an amino acid alteration at Arg 65. The relatedness of this intermediate to an uncleaved proalbumin variant containing a glutamine for arginine substitution²² emphasizes the essential pairing of basic residues for substrate recognition of both peptide and protein precursors by processing enzymes.

Our analysis of the variant proinsulin intermediate des-Arg 31, Arg 32-[X 65] proinsulin in patients with familial hyperproinsulinaemia identifies a new locus of mutation within the insulin gene. These loci include sites in the B-chain region¹⁻⁴, the connecting sequence between the C peptide and the insulin A chain (this report), the 3'-untranslated region^{5,6} and the two intervening sequences⁵. Nevertheless, the structural and physiological consequences of these changes vary considerably: the first results in a normally processed, but functionally abnormal hormone; the second results in an abnormally processed intermediate of conversion (with the likely potential for conversion to the normal hormone); the third and fourth probably result in completely normal hormone products. Another proinsulin intermediate, previously isolated from a different family with hyperproinsulinaemia, and reported to have the C peptide joined to the insulin B chain⁸, has not been analysed in detail. The complexity of both post-transcriptional^{5,6} and post-translational^{11-13,23,24} processing in insulin biosynthesis suggests still other sites (especially those for excision and ligation of precursor mRNA and for cleavage of the signal peptide from pre-proinsulin) at which mutations would yield abnormal gene products.

This work was supported by grants AM 18347, AM 13941 and AM 20595 from the National Institute of Arthritis, Metabolism, and Digestive Diseases of the NIH. D.C.R. is the recipient of National Science Research Award AM 05723 from the NIH. H.S.T. is the recipient of Research Career Development Award AM 00145 from the NIH.

Received 12 January; accepted 17 March 1981.

1. Tager, H. *et al.* *Nature* **281**, 122-125 (1979).
2. Given, B. D. *et al.* *New Engl. J. Med.* **302**, 129-135 (1980).
3. Olefsky, J. M., Saekow, M., Tager, H. & Rubenstein, A. H. *J. biol. Chem.* **255**, 6098-6105 (1980).
4. Tager, H. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **77**, 2410-2414 (1980).
5. Bell, G. I. *Nature* **284**, 26-32 (1980).
6. Ullrich, A., Dull, T. J., Gray, A., Brosius, J. & Sires, I. *Science* **209**, 612-615 (1980).
7. Gabbay, K. H., DeLuca, K., Fischer, J. N., Mako, M. E. & Rubenstein, A. H. *New Engl. J. Med.* **294**, 911-915 (1976).
8. Gabbay, K. H., Bergenstal, R. M., Wolff, J., Mako, M. E. & Rubenstein, A. H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2881-2885 (1979).
9. Kanazawa, Y., Hayashi, M., Ikeuchi, M., Hiramatsu, K. & Kosaka, K. *Eur. J. clin. Invest.* **8**, 327-331 (1978).
10. Kanazawa, Y. *et al.* *Proc. Symp. Proinsulin, Insulin and C-Peptide*, Tokushima (eds Baba, S., Kaneko, T. & Yanihara, N.) 262-269 (Excerpta Medica, Amsterdam, 1979).
11. Steiner, D. F. *Recent Prog. Horm. Res.* **24**, 207-282 (1969).
12. Chance, R. E. in *Diabetes* (eds Rodriguez, R. R. & Vallance-Owen, J.) 292-305 (Excerpta Medica, Amsterdam, 1971).
13. Steiner, D. F. *Diabetes* **26**, 322-340 (1977).
14. Horwitz, D. L., Starr, J. I., Mako, M. E., Blachard, W. G. & Rubenstein, A. H. *J. clin. Invest.* **55**, 1278-1283 (1975).
15. Cuatrecasas, P. *J. biol. Chem.* **245**, 3059-3065 (1970).
16. Starr, J. I., Horwitz, D. L., Rubenstein, A. H. & Mako, M. E. *Methods of Hormone Radioimmunoassay* (eds Jaffe, B. M. & Behrman, H. L.) 613-629 (Academic, New York, 1979).
17. Varandani, P. T. *Biochemistry* **6**, 100-104 (1967).
18. Kuzuya, H., Blix, P. M., Horwitz, D. L., Steiner, D. F. & Rubenstein, A. H. *Diabetes* **26**, 22-29 (1977).
19. Faber, O. K. *et al.* *Diabetes* **27**, Suppl. 1, 170-177 (1978).
20. Cole, R. D. *Met. Enzym.* **11**, 206-208 (1967).
21. Davis, B. J. *Ann. N. Y. Acad. Sci.* **121**, 404-422 (1964).
22. Brennan, S. O. & Carrell, R. W. *Nature* **274**, 908-909 (1978).
23. Chan, S. J., Keim, P. & Steiner, D. F. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1964-1968 (1976).
24. Tager, H. S. *et al.* *Ann. N. Y. Acad. Sci.* **343**, 133-147 (1980).

Allosteric regulation of crocodilian haemoglobin

M. F. Perutz*, C. Bauer†, G. Gros‡, F. Leclercq§, C. Vandecasserie§, A. G. Schnek§, G. Braunitzer||, A. E. Friday¶ & K. A. Joysey¶

* MRC Laboratory of Molecular Biology, Cambridge CB2 2QH, UK

† Institut für Physiologie, Universität, 8400 Regensburg, FRG

‡ Physiologisches Institut, Universitätsklinikum 4300 Essen, FRG

§ Chimie General I, Université Libre, Bruxelles, Belgium

|| Max-Planck Institut für Biochemie, 8033 Martinsried, FRG

¶ University Museum of Zoology, Cambridge CB2 3EJ, UK

The oxygen affinity of most vertebrate haemoglobins in the absence of diffusible electrolytes is much higher than that of blood. In the red cell this affinity is lowered by organic phosphates, hydrogen ions, chloride ions and CO₂ (refs 1–5). Similarly, crocodilian haemoglobin also has a much higher oxygen affinity than crocodile blood, but this is due to bicarbonate ions⁶, as neither phosphates nor carbamino CO₂ lower its oxygen affinity, and chloride does so only weakly⁷. The complete sequences of the haemoglobins of the caiman, the Nile crocodile and the Mississippi alligator (to be reported elsewhere⁸) show 102 substitutions between human and caiman, and 123 between human and the other two crocodilian haemoglobins. Here we consider how these substitutions may explain the changes in allosteric control, and also their bearing on the phylogenetic relationships between the crocodilians and other groups of bony vertebrates. We propose that a few of the substitutions abolish or weaken the binding sites for the usual allosteric effectors and create a new pair of binding sites which are complementary to bicarbonate ions in the deoxy (T) structure, but not in the oxy (R) structure.

Consider first the loss of allosteric inhibition by organic phosphates, carbamino CO₂ and chloride (Table 1). In human deoxyhaemoglobin, organic phosphates are bound by the four pairs of cationic residues that lie in the cavity between the two β -chains: Val NA1 (1), His NA2 (2), Lys EF6 (82) and His H21 (143)⁹. In the three crocodilian haemoglobins, the histidines are replaced by neutral residues. In two of the haemoglobins the α -amino groups are acetylated; in the third, a proline in position 2 forces a bend in the chain which places the α -amino groups out of reach of the phosphates. This leaves only the two lysines which are known from human fetal haemoglobins F₁ to be insufficient for phosphate binding, although His NA2 is still present there¹⁰. In human deoxyhaemoglobin, carbamino CO₂

bound to Val 1 α accepts a hydrogen bond from Ser H14 (131) α (ref. 11). In the three crocodilian haemoglobins this is replaced by Ala. This replacement may also inactivate one of the strong chloride binding sites which probably coincides with the CO₂ site¹². In human deoxyhaemoglobin, another pair of CO₂ molecules competes with organic phosphates for binding to Val 1 β ; when bound, each carbamino CO₂ is stabilized by a salt bridge to Lys 82 β (ref. 11). In caiman haemoglobin, the bend in the chain forced by Pro 2 inhibits that bridge; in the other two crocodilian haemoglobins, the blocking of Val 1 β inhibits binding of CO₂.

Two equivalents of bicarbonate ion per tetramer are bound to caiman deoxyhaemoglobin with an association constant of $2.0 \times 10^{-3} \text{ M}^{-1}$, whereas oxyhaemoglobin shows no significant binding. In the presence of CO₂, discharge of oxygen produces no change in pH, because the protons taken up by haemoglobin are neutralized by those set free in the formation and binding of bicarbonate⁶. At pH 7.4 and 20 °C in 0.1 M NaCl, application to 40 torr CO₂ raises P_{50} ninefold from 3.4 to 30.0 torr; this is comparable to the 10-fold rise, from 4.35 to 44.0 torr, produced by the addition of 2 mM inositolhexaphosphate (IHP) to a similar solution of human haemoglobin A at 25 °C (ref. 13). In human haemoglobin A, IHP reduces Hill's coefficient from 2.9 to 2.4; in crocodile haemoglobin, CO₂ reduces it to the same value. In both cases the allosteric equilibrium is shifted too far towards the T structure for oxygen binding to be fully cooperative. At pH 7.0, addition of CO₂ to caiman methaemoglobin produces a large negative circular dichroism peak at 287 nm, diagnostic of the quaternary T structure; the same effect is produced by IHP in human methaemoglobin A^{14,15}.

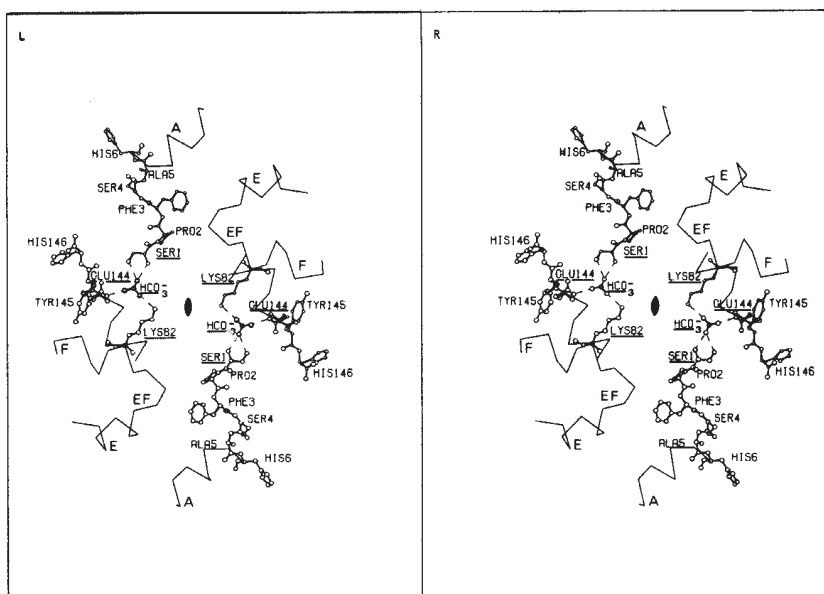
We have examined all possible sites in the model of human deoxyhaemoglobin where the amino acid replacements indicated by the three crocodilian sequences could possibly cause oxygen-linked binding of bicarbonate ions, and have found only one pair of sites with the right stereochemistry. These lie in the cavity between the two β -chains where organic phosphates or carbamino CO₂ are bound in other species, and are formed by Lys EF6(82) and Glu HCl(144) of one β -chain together with the N-terminal residue of its partner chain (Fig. 1, Table 1). In caiman haemoglobin, the N-terminal sequence is Ser-Pro-Phe-. A model of the bicarbonate binding site was first built manually by replacing the side chains in human deoxyhaemoglobin with those of caiman, and was then refined by computer. The main chain from the α -carbon of Pro 2 onwards was held in the same position as in human deoxyhaemoglobin. At this atom the chain is forced to turn a corner at an angle defined by the stereochemistry of the proline. This rigid turn brings the N-terminal serine within exact reach of the bicarbonate ion, so that one of the bicarbonate oxygens can form a salt bridge with the α -NH₃⁺ and can also accept a rather long hydrogen bond from the serine OH. The second bicarbonate

Table 1 Amino acid replacements for allosteric control in crocodile haemoglobin

Position in			Species					Functions
Structure	Sequence		Human	Other bony vertebrates	Caiman	Nile crocodile	Mississippi crocodile	
NA1	β	1	Val	Val or Gly	Ser	Ac-Ala	Ac-Ala	In human and many other vertebrate deoxy-Hbs, the α -amino groups bind either organic phosphate or carbamino CO ₂ , which in turn forms a salt bridge with Lys 82 of the same β -chain. In crocodile Hb, blocking of α -NH ₃ ⁺ or Pro in position 2 inhibits these functions
NA2	β	2	His	Gln, Asn, Glu, His, Asp, Met	Pro	Ser	Ser	In human Hb, His, and in many other vertebrates Gln or Asn, bind organic phosphates
EF6	β	82	Lys	Lys	Lys	Lys	Lys	Lys binds diffusible anions in the internal cavity of deoxy-Hbs. Invariant in all vertebrate Hbs except trout I which lacks all oxygen-linked functions and where it is replaced by Leu*
H21	β	143	His	Lys, Arg, Ser, His	Ala	Ala	Ala	His, Lys or Arg form salt bridges with organic phosphates
HCl	β	144	Lys	Arg, Ala, Ser, Gln, Lys	Glu	Glu	Glu	In human Hb, Lys is free and the same is probably true of Arg, Ser or Gln in other species. Glu can accept an H-bond from HCO ₃ ⁻ whose charge is compensated by Lys 82
H14	α	131	Ser	Very variable	Ala	Ala	Ala	In human Hb, OH of Ser forms an H-bond with carbamino CO ₂ or Cl ⁻ bound to Val 1 α , but this bond is absent in many species

* Unpublished data (D. Barra, F. Bossa and M. Brunori).

Fig. 1 Stereo drawing of proposed bicarbonate binding site between the two β -chains of caiman deoxyhaemoglobin. The central sign marks the dyad symmetry axis. Bicarbonates and their binding residues are underlined. Capital letters mark helical and inter-helical segments. The drawing was prepared using the PLUTO program and Fermi's refined atomic coordinates of human deoxyhaemoglobin¹⁰. Certain of the amino acid residues in the atomic model of this haemoglobin were replaced by those in caiman; the bicarbonate ions were attached to them and all the new atomic positions were then measured manually. These rough coordinates were then corrected to standard bond distances and angles using Diamond's BILDER program²¹ on the Evans and Sutherland Picture System I linked to a DEC PDP 11/50 computer. The hydrogen bond distances quoted in the text were derived from this computer-generated model. The hydrogen bond from HCO_3^- to COO^- of Glu 144 looks straight in the picture, but in fact the two groups lie at different heights, making the bond angle $\text{COH } 110^\circ$, as it should be.



oxygen can form a salt bridge with Lys EF6(82) and the third oxygen can donate a hydrogen bond to one of the carboxylate oxygens of Glu HC1(144). In this model all hydrogen bonds come out reasonably: Ser $\text{N}-\text{O}_1 = 2.8 \text{ \AA}$; Ser $\text{O}_\gamma-\text{O}_1 = 3.3 \text{ \AA}$; Lys $\text{N}_\epsilon-\text{O}_2 = 2.9 \text{ \AA}$; Glu $\text{O}_\alpha-\text{O}_3 = 2.7 \text{ \AA}$. All side chains are in the strain-free, straight, staggered conformation. Even so, the model should be checked by X-ray analysis. In the haemoglobins of the Nile crocodile and the Mississippi alligator, where the N-terminal sequence is acetyl-Ala-Ser-Phe, the α -NH of acetylalanine could donate a hydrogen bond to O_1 of the bicarbonate at the same distance as the α -NH $_2$ does in caiman, provided the chain turns a corner at Ser 2 at the same angle as it does at Pro 2 in caiman. Again, this should be checked by X-ray analysis.

The salt bridges and hydrogen bonds made by the two symmetry-related bicarbonate ions would stabilize the relative positions taken up by the two β -chains in the quaternary T structure; furthermore, by immobilizing Glu HC1(144), they would help to clamp the three C-terminal residues Glu-Tyr-His in their rigid deoxy conformation, with the tyrosine hydrogen-bonded to Val FG5 β and the histidine salt-bridged to Asp FG1 β and Lys C5 α . On transition to the R structure, the two β -chains move relative to one another by several Ångströms and their C-terminal residues become mobile, so that the bicarbonate binding sites are destroyed.

Other small anions also lower the oxygen affinity, but much less than bicarbonate. When the concentrations of chloride, nitrate, perchlorate or phosphate are increased from 0.01 M to 0.5 M at pH 7.0 and 25 °C, P_{50} is lowered by factors of 2.0, 1.6, 1.7 and 1.3, respectively, whereas the same increase in the concentration of bicarbonate lowers P_{50} 13-fold⁶. In human haemoglobin, chloride probably binds to Lys 82 β (ref. 16). If the same is true of caiman haemoglobin, the binding of chloride and

bicarbonate should be competitive; in fact, a rise in $[\text{Cl}^-]$ from 0.01 to 0.5 M approximately halves the effect of HCO_3^- on P_{50} . NO_3^- is a planar ion like HCO_3^- , and the NO distances are only 0.1 Å shorter than those of CO. NO_3^- could therefore bind to Ser 1 and Lys 82, but it has no hydrogen to donate to Glu 144; nor does ClO_4^- . HPO_4^{2-} could do so, but because this ion is tetrahedral, binding of one oxygen to Ser 1 and another to Lys 82 makes the OH bond of the third oxygen point away from the carboxylate of Glu 144, so that no hydrogen bond can be formed.

The decrease in oxygen affinity brought about by the interaction of crocodilian haemoglobin with bicarbonate ensures that oxygen is released from the blood to the tissues at a relatively high partial pressure of oxygen. If the haemoglobin were insensitive to bicarbonate, the venous P_{O_2} would be only 7 torr; interaction with bicarbonate raises this to 27 torr, thus creating a large enough pressure head for the flow of oxygen from the blood to the tissues¹⁷. In human haemoglobin, the same relative rise in P_{O_2} , from 19 to 40 torr, can be brought about only by the combined effects of CO_2 and diphosphoglycerate. It is surprising that the simple and direct reciprocating action between oxygen and one of the end products of oxidative metabolism has not been adopted by other vertebrates. We do not know what advantage it gives to the crocodilians.

Our results show that an entirely new function can evolve in a protein by no more than three amino acid substitutions, requiring only four nucleotide base changes; just two more amino acid substitutions, or three base changes, are needed for inhibition of the old functions (oxygen-linked phosphate, carbamino CO_2 and Cl^- binding). Most of the other ~100 substitutions that distinguish crocodilian from human haemoglobins are conservative and would have little if any effect on the oxygen equilibrium. There are some significant substitutions in the haem

Table 2 Number of amino acid substitutions between the α -globin chains of different species

	Man	Caiman	Mississippi alligator	Nile crocodile	Goldfish	Trout	Tadpole	Chicken	Goose	Viper
1. Caiman	46									
2. Mississippi alligator	47	20								
3. Nile crocodile	44	21	17							
4. Goldfish	65	68	68	64						
5. Trout	60	64	64	66	43					
6. Tadpole	61	65	67	66	70	63				
7. Chicken	35	46	49	45	72	70	70			
8. Goose	41	39	35	30	64	61	64	30		
9. Viper	50	64	66	62	76	70	69	57	59	
10. Kangaroo	27	43	43	43	67	60	60	41	39	54

1, *Caiman crocodilus*; 2, *Alligator mississippiensis*; 3, *Crocodylus niloticus*; 4, *Carassius auratus*; 5, *Salmo irideus*; 6, *Rana catesbeiana*; 7, *Anser anser*; 8, *Vipera aspis*; 9, *Vipera aspis*; 10, *Macropus giganteus*.

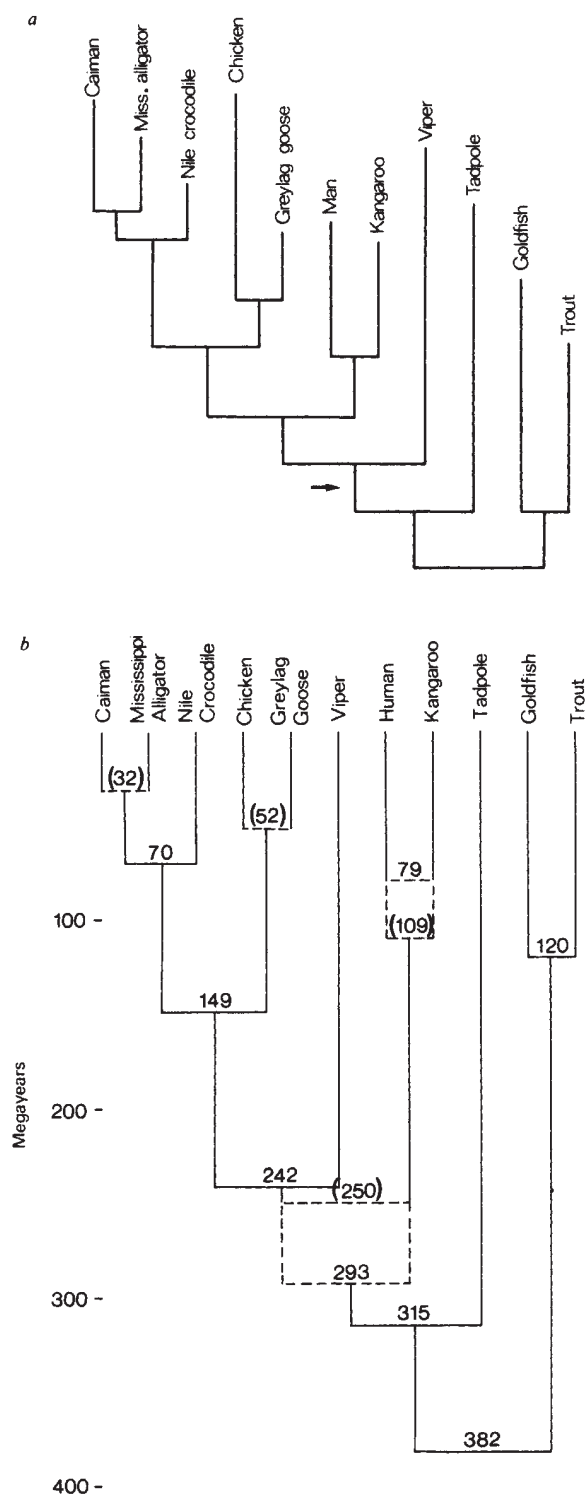


Fig. 2 *a*, Wagner network computed from Table 2 by the method of Farris²². The root has been placed midway along the branch separating the two fish from the other species, in accord with conventional zoological opinion. The arrow shows an alternative position of the root at the midpoint between the two species separated by the largest network distance. This network has a total length of 249.5; an alternative network, with the shortest length of 249.0, has a much poorer fit to the observed differences. *b*, Phylogenetic tree based on comparative anatomy and the fossil record. The number at each branching point is an estimate of the minimum date of divergence of the ancestors of the living forms, based on direct fossil evidence. Alternative possible dates, indicated by broken lines and numbers in parentheses, refer to less reliable dates. The justification for some of these dates is given in ref. 19 and the method of dating is further discussed in ref. 23.

pockets and at the subunit contacts, but none of these is unique or could play any part in the allosteric control by bicarbonate ions. All the residues which are essential for the formation of the characteristic T and R structures are either conserved or have been replaced by ones that can serve the same purpose equally well.

Do the crocodilian sequences tell us anything interesting about the phylogenetic relationships between these reptiles and other groups of bony vertebrates? The α -globin sequences of 21 mammals, 3 birds, 3 fish, 1 amphibian and 1 other reptile (viper) are now known, but not the β -globin sequence of the viper. We have selected two species from among each of the mammals, birds and fish and compared their α -globin sequences with those of the crocodilians, the bullfrog tadpole and the viper. Table 2 shows a matrix of amino acid differences between all possible pairs of sequences; the data in Fig. 2*a* have been calculated from this matrix by an algorithm that does not allow the difference between any pair of sequences to be less than the observed one; it attempts to minimize the total length of the network and to maximize its fit to the observed differences. It shows that the crocodilian sequences are most similar to those of the birds, consistent with fossil evidence suggesting that birds arose from archosaurian reptiles. As might be expected, the number of amino acid differences and the network distances between each of the crocodilians and the viper (a lepidosaur reptile) are much larger than between the crocodilians and the birds, but it is surprising that they are also much larger than between the crocodilians and the mammals. Figure 2*b* shows, for comparison, a zoologically conventional phylogenetic tree, derived from the evidence of comparative anatomy and the fossil record, which shows the crocodilians as more closely related to the viper than the mammals. For the rest, Fig. 2*a* and *b* are similar in pattern, although Fig. 2*b* has very unequal pairs of branches arising from the same node. These support the conclusions of Romero-Herrera *et al.*^{18,19}, based on myoglobin, that the rates of change differ on different branches of the phylogenetic tree. Without knowing more reptile sequences we cannot be sure, therefore, whether the large differences between the crocodilians and the viper suggest a divergence between the ancestors of archosaurian and lepidosaurian reptiles earlier than that between the archosaurian reptiles and mammals, or are due merely to accelerated fixation of mutations along one particular evolutionary lineage. This question is worth following up.

We thank Mr A. Stangl and Mrs B. Schrank for their help in determining the amino acid sequences, Professors K. Johansen and R. Coulson for gifts of crocodile blood, Drs S. E. V. Phillips and B. Shaanan for their help with Fig. 1 and Professor Howard Morris for checking an N-terminal sequence by mass spectrometry.

Received 24 March; accepted 8 May 1981.

- Baldwin, J. M. *Prog. Biophys. molec. Biol.* **29**, 225-320 (1975).
- Perutz, M. F. *Br. med. Bull.* **32**, 191-208 (1976).
- Kilmartin, J. V. *et al. Biochim. biophys. Acta* **534**, 15-25 (1978).
- Kilmartin, J. V. & Rossi-Bernardi, L. *Physiol. Rev.* **53**, 836-890 (1973).
- Kilmartin, J. V. *Trends biochem. Sci.* **2**, 247-250 (1977).
- Bauer, C. *et al. J. biol. Chem.* (in the press).
- Bauer, C. & Jelkmann, W. *Nature* **269**, 825-827 (1977).
- Leclercq, F., Schnek, A. G., Braunitzer, G., Stangl, A. & Schrank, B. *Hoppe-Seyler's Z. physiol. Chem.* (in the press).
- Arnone, A. *Nature* **237**, 146-149 (1972).
- Bunn, H. F. & Briehl, R. W. *J. clin. Invest.* **49**, 1088-1095 (1970).
- Arnone, A., Rogers, P. H. & Briley, P. D. in *Biophysics and Physiology of Carbon Dioxide* (eds Bauer, C., Gros, G. & Bartels, H.) 67-74 (Springer, Berlin, 1980).
- O'Donnell, S., Mandaro, R., Schuster, T. M. & Arnone, A. *J. biol. Chem.* **254**, 12204-12208 (1979).
- Imai, K. *J. biol. Chem.* **249**, 7607-7612 (1974).
- Perutz, M. F., Fersht, A. R., Simon, S. R. & Roberts, G. C. K. *Biochemistry* **13**, 2174-2186 (1974).
- Fermi, G. & Perutz, M. F. *J. molec. Biol.* **114**, 421-431 (1977).
- Perutz, M. F. *et al. J. molec. Biol.* **138**, 649-670 (1980).
- Jelkmann, W. & Bauer, C. *Comp. Biochem. Physiol.* **A65**, 331-336 (1980).
- Romero-Herrera, A. E., Lehmann, H., Joysey, K. A. & Friday, A. E. *Nature* **246**, 389-395 (1973).
- Romero-Herrera, A. E., Lehmann, H., Joysey, K. A. & Friday, A. E. *Phil. Trans. R. Soc. B283*, 61-163 (1978).
- Fermi, G. *J. molec. Biol.* **97**, 237-256 (1975).
- R. Diamond in *Biomolecular Structure, Conformation and Evolution* Vol. 1 (ed. Srinivasan, R.) 567-588 (Pergamon, Oxford, 1980).
- Farris, J. S. *Am. Nat.* **106**, 645-668 (1972).
- Joysey, K. A. *Symp. zool. Soc. Lond.* **46** (in the press).

BOOK REVIEWS

Reason and unreason in assessing risk

Eric Ashby

LAST November the Royal Society held a discussion meeting on the assessment and perception of risk. Among the speakers were engineers, toxicologists, psychologists, mathematicians, biologists, administrators; but there was no representative of the profession which makes its living out of risk-assessment: the actuaries who calculate insurance premiums. If the experts in this profession are not cautious enough, the firm loses money. If they are too cautious, the firm loses business. The impression left (upon me, at any rate) by this discussion was that research on risk-assessment, though much of it is interesting and sophisticated, is not yet a reliable tool for people who have to make decisions in the public interest about hazardous projects. Statisticians can give accurate estimates of familiar risks. Social psychologists can give plausible descriptions of the way these risks are perceived. What we need is more evidence from the people who may lose their reputations or their jobs if they misjudge risks. With this in mind I welcomed the invitation to read two more books on this subject.

The book edited by Conrad is a record of a workshop organized by the Battelle Institute. There are 20 papers by authors from different countries. Part I covers familiar ground; it deals with the methods used in risk-analysis. Nothing fresh here, except a perceptive criticism of the "revealed preference theory" popularized by Starr in a paper in *Science* 12 years ago (165, 1232-1238). Parts II, III and IV cover the social and political consequences of risk-assessments. Here there is a good deal of interesting and controversial opinion.

Risk-assessment is only one of several operations necessary for the much broader discipline of decision-analysis, and it is a pity that it has been split off and isolated from the parent subject. For when you have made your statistical assessment (of, for instance, lives that would be saved by the wearing of seat belts in cars), and your guess as to how motorists perceive this risk, you still lack data essential for making a decision. The politician wants an answer to the question: "What is an acceptable risk?". But there is an antecedent question to that one: "Who, in a pluralistic democracy, has the right to decide what is an acceptable risk?". When the issue arises in Parliament, this is the sort of question which MPs ask. And what is the purpose behind a statement about risk-assessment? Is it to impose safety standards for which there is a public demand? Is it to reassure the public that some decision already

Society, Technology and Risk Assessment. Edited by J. Conrad. Pp.304. ISBN 0-12-186450-2. (Academic: 1980.) £15.20, \$35. *Dealing with Risk: The Planning, Management and Acceptability of Technological Risk*. Edited by R.F. Griffiths. Pp.144. ISBN 0-7190-0819-0/0-470-27136-1. (Manchester University Press/Halsted: 1981.) £11.50, \$19.95.

taken, for reasons nothing to do with safety, need not cause disquiet? Then there is the assumption, one I have myself made, that public perceptions of risk are irrational (for example the person who willingly indulges in some practice which exposes him to a 10^{-3} risk of death but who campaigns fanatically against some other practice which would expose him to a 10^{-7} risk of death). Irrational, no doubt, on the criteria applied professionally by scientists, but most of the British public are not scientists and their irrationalities are as significant for the decision-analysis as are the hard data from the statisticians. To exclude their views from the risk-benefit balance sheet would be to cook the books. Social intuition may appear to be a woolly concept compared with fault-tree-analysis, but it is likely to carry more weight in Parliament and in Whitehall.

Owing to considerations such as these, a good deal of research on risk-assessment seems peripheral to the issues that have to be decided by developers or planning authorities. This is the opinion of some of the contributors to the Battelle Institute workshop. What emerges from these papers is the impression that (in the words of Volker Ronge, from Germany)

... decision-making processes have apparently come out to be risky in that they are no longer capable of producing sufficient social consent to outcomes that have to a large amount been "decided upon" outside the political realm.

The book edited by Conrad, then, enlarges our awareness (if not our understanding) of the sociological and political ingredients in decision-analysis. The book edited by Griffiths is a more pragmatic set of papers, most of them written by people who really have to make decisions about risky projects and who work at the coal-face of public opinion. It is therefore refreshingly practical. A safety adviser from industry sets out clearly how he has to compose his advice. No use to pretend that there is zero risk. Sooner or later some accident will happen and there will be fatalities. There is a temptation to rely on compared risks ("we should not spend money on raising the height of the

handrails if the risk of falling over them is smaller than the risk of being harmed by toxic chemicals in the atmosphere"). But of course this sort of argument is not acceptable to the industry or to the public. The decision-maker is driven back to crude cost-benefit analysis: the cost of saving a statistical life. Thus (an example from Conrad's book) in the Netherlands it was considered whether to introduce earth leakage switches into domestic wiring systems. It would reduce the number of electrocutions by a factor of ten; but since the present risk is only 5×10^{-7} per year, the cost per life saved would be of the order of \$2 millions.

It is no wonder that this sort of arithmetic is not commonly disclosed to the public in planning enquiries, nor, if it were, would it make much impression. In two chapters on hazard and risk in planning, two planners (one from the Cheshire County Council and the other from the Halton Borough Council in Cheshire) explain exactly what procedures they follow when there are hazards in the project being planned. The book is worth reading for these two contributions alone; for the experience of these two practical men may do something to remedy the most serious problem in risk-assessment, namely the alienation of the public toward those who have to decide, on the public's behalf, what is an acceptable risk. Risk, as Stephen Cotgrove points out in the concluding essay in *Dealing with Risk*, "is not just a statistical calculation. It is also a moral judgement about defensible conduct". It is the legitimacy of decisions made about risk, more than their accuracy, which is being called in question by public interest groups and which causes exasperating delays in what are often desirable developments. The open, empirical and unsophisticated procedures described by the planners from Cheshire, coupled with their willingness to take the public into their confidence, may do more to restore the legitimacy of planners than would volumes of research on risk-assessment.

These two books, together with the book *Societal Risk Assessment* (reviewed in *Nature* 289, 827; 1981) should dispel any illusion that the problem: How safe is safe enough? is within sight of solution. But why, oh why, have we not yet heard evidence from actuaries who calculate insurance?

Lord Ashby is Chancellor of Queen's University, Belfast, and a Fellow of Clare College, Cambridge.

Ninth stage in the long march of Joseph Needham

R.G.W. Anderson

Science & Civilisation in China, Vol. V:4. By Joseph Needham. Pp.772. ISBN 0-521-08573-X. (Cambridge University Press: 1980.) £48, \$105.

DR Needham's great compilation at this stage occupies a substantial length of bookshelf — already more, for example, than Thorndike's monumental *A History of Magical and Experimental Science*. Though this volume is the ninth to appear, a recently issued "State of the Project" pamphlet states that the completed series will comprise 20 books. There will be three further volumes on chemistry and chemical technology after the currently reviewed volume, the third to appear.

Three major areas are covered by this book: chemical techniques, alchemical theory and comparative macrobiotics. The first is largely involved with a description of distillation processes. Here, as with many other topics treated by Needham in the series, the reader is not simply provided with an elucidation of the Chinese tradition, but with the best comparative account available for a number of cultures. However one main argument has now been shown to be inaccurate: the suggested greater efficiency of the Eastern stills to the simple Western types (retort or alembic). Needham proposes that the relatively late date for the distillation of alcoholic spirits in the West (13th century AD) can be explained by the ineffectiveness of the condensation surface of the retort and alembic. In China, he maintains, the more efficient stills account for evidence of spiritous liquors in the 7th century AD. This argument has been demolished in a recent paper in *Ambix* (27, 69; 1980) (of which Needham, it must be admitted, is co-author) in which stylized Western retorts, Mongolian stills and Chinese stills were tested experimentally and were shown to be equally efficient at alcohol distillation. How odd, then, that there is no positive evidence of earlier distillation of spirits in the West as there is of perfume oils and petroleum fractions.

Much less is said about other processes and the associated apparatus, though there is an interesting discussion of reaction-chambers (which do not have an exact Western counterpart). Cunningly designed water-cooled examples are illustrated, redrawn from the *Chin Hua Chung Pi Tan Ching Pi Chih* ("Confidential Instructions on the Manual of the Heaven-piercing Golden Flower Elixir") of 1225 AD. So bizarre are some of these that one wonders if they were really constructed.

The second section, "Alchemical Theory", discusses the application of Taoist philosophy to an understanding of chemical materials and their reactions. To the Western mind of today, the explanations of the Chinese alchemists bear little

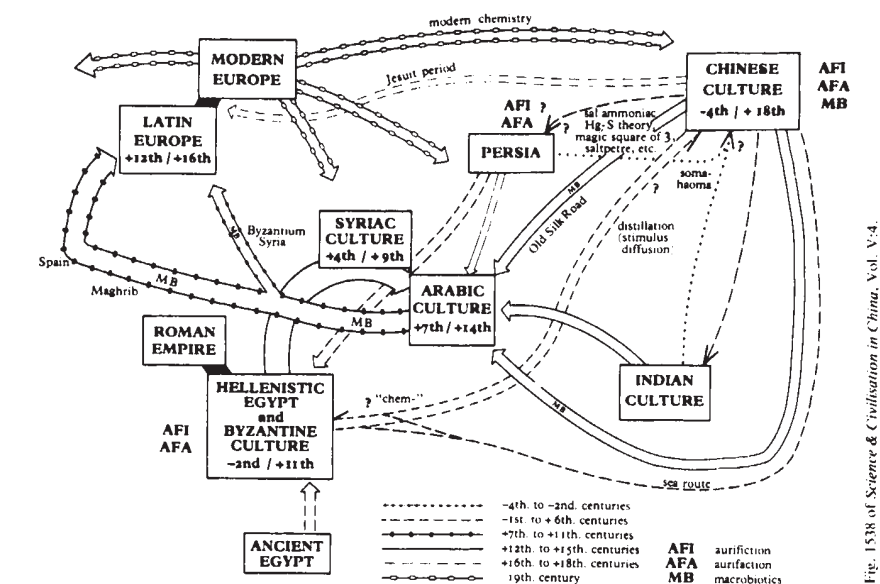


Chart illustrating the multifocal origins of proto-chemistry, to be seen as superimposed roughly upon a map of the Old World. See p.504 of the book for further details.

or no relationship to current scientific understanding; Needham himself says that "the dominant goal of proto-scientific alchemy was contemplative" (p.244) — though this must be reconciled with a later statement that theoretical alchemy was "a deductive proto-science on quite the same level as medicine, acoustics or magnetic geomancy" (p.298).

In the final part of the book, macrobiotics is defined as longevity or material immortality attained with the aid of chemical knowledge. Here a comparison is made of Chinese alchemy, Hellenistic proto-chemistry (macrobiotics played little part in Graeco-Egyptian culture) and Arabic and Latin alchemy. Most importantly, the interactions and hence influences are probed. Thus it is shown that the early chemistry of the Arabs was not exclusively Graeco-Egyptian in origin, but included contributions from China, India and Persia. A summing-up is provided in the form of a chart (reproduced above), illustrating the complex multifocal origins of proto-chemistry.

The reader of this authoritative and rewarding book is left with the impression of two themes pursued relentlessly by Needham. First, routes of influence are constantly sought, the chances of independent discovery of the same solution of the same problem in geographically distant places only being considered when all else fails. This can produce some rather absurd, tongue-in-cheek suggestions such as the relationship between the Mongolian still and the Irish poteen still. Second, is the effort made to claim primacy for the Chinese on every occasion. Thus in a passage attempting to pre-date Western spirit distillation by the Chinese, Needham follows some rather ambiguous early

evidence with the revealing remark "We still have plenty of time before us ere we reach the +12th century, time of the first distillation in the West".

Yet it is the highly personal approach which helps to make this prodigious work so enjoyable. How delightful and revealing, for example, is the seemingly careless comment: "Resting one day in 1968 at a wayside auberge in the South of France, I read again the text of the *Tabula Smaragdina* and felt so much the Chinese flavour of it that I started translating it to see how it would look in that language".

R.G.W. Anderson is Keeper of the Department of Chemistry at the Science Museum, London.

Bright future of NSE

B. E. F. Fender

Neutron Spin Echo. Lecture Notes in Physics, 128. Edited by F. Mezei. Pp.253. ISBN 3-540-10004-0. (Springer-Verlag: 1980.) \$19.50, DM 33.

NEUTRON spin echo (NSE) techniques represent the single most important innovation in neutron scattering methods in recent years. The idea originated with F. Mezei and has been skilfully realized at the Institut Laue-Langevin in a full working instrument by J. Hayter and Mezei.

Conventional neutron inelastic scattering instruments for studying the dynamics of materials are conceptually simple; neutrons of a particular energy are selected by a monochromating crystal or a

velocity selector and, after interaction with a sample, the neutron energy is redetermined. The instrument design is dominated by the relatively low flux of neutron beams and this problem is particularly acute when a high-energy resolution is required, for example to probe slow motions in polymer chains. The breakthrough which the spin echo technique represents is that the velocity of the neutron is determined from the Larmor precession of the neutron spin in a magnetic field. In this way the resolution of the energy change of scattering does not depend on the monochromatization of ingoing or outgoing neutron beams, and very small energy changes down to 10^{-8} eV can be measured.

This book, which is based on a workshop at the ILL, is in three parts. The first deals with the method from the point of view of both theory and its practical realization. The discussion is sufficiently detailed for the specialist to appreciate as a review, while at the same time it provides a useful introduction for the intending practitioner.

The non-specialist will probably be more

interested in the second part which gives a range of examples. Even though most of the measurements are preliminary, this section well illustrates how improvements in neutron instrumentation generate fresh scientific opportunities. The chapters which are likely to be the best indicators of future research are those on the dynamics of micelle solutions, polymer solutions, spin glass dynamics and superfluid ^4He , although the reader must beware of the rapidity of developments in these and other fields. The third section deals primarily with future instrumentation; it will serve for some time as food for thought for the experimenters of tomorrow. Apart from concluding comments by the editor, there are also interesting chapters on the use of spin echo methods to discriminate against thermal diffuse scattering and on the application of NSE techniques in order to measure phonon line widths.

In sum, then, this is a thoroughly worthwhile contribution to a good series.

B. E. F. Fender is Directeur Adjoint at the Institut Max von Laue-Paul Langevin, Grenoble.

Reluctant leaf-eaters of Barro Colorado

R.D. Martin

The Foraging Strategy of Howler Monkeys: A Study in Primate Economics. By Katharine Milton. Pp.165. ISBN 0-231-04850-5. (Columbia University Press: 1980.) \$20, £13.

IT IS generally agreed that C.R. Carpenter's 1934 monograph (*Comp. Psychol. Monogr.* 10, 1-168) on mantled howler monkeys, *Alouatta palliata*, marked a significant transition in primate field studies, from essentially anecdotal accounts to serious scientific investigations. His study was conducted on the small island of Barro Colorado (BCI), an area of about 1,500 hectares of rainforest isolated during the construction of the Panama Canal in 1914. The Smithsonian Tropical Research Institute has, over a period of several decades, sponsored numerous fundamental studies of the fauna and flora. In the half century since Carpenter's publication, quite a few studies have been conducted on the primate species inhabiting BCI, and the howler monkey population (now numbering over 1,000 individuals) has been re-studied and re-censused at fairly regular intervals. *Alouatta palliata* therefore counts among the best-known primate species with respect to natural behaviour and ecology.

The first systematic primate field studies were, inevitably, largely concerned with basic data collection and primarily qualitative. As work has progressed, there has

been increasing emphasis on quantification of both behavioural and ecological parameters, and most modern studies are specifically designed to answer well-defined questions. There has also been increasing integration of laboratory studies with field work as more specific issues, such as nutritional and toxic properties of natural foodstuffs, have come into focus. All of these modern trends are admirably exemplified in Katharine Milton's *The Foraging Strategy of Howler Monkeys*, which provides a fine example of the great advances made since Carpenter's pioneering work. Appropriately, Milton's main 14-month field study was also conducted on BCI, on the direct descendants of the monkeys studied by Carpenter.

The howler is one of the largest of the New World monkeys (Ceboidea) and has been widely regarded as the most folivorous. Milton's quantitative, year-round data confirm that, on average, the mantled howler spends more time eating leaves than either fruits or flowers. Yet the howler lacks any obvious specialization of its digestive tract to extract particular benefit from an abundant gut flora in processing the leaves in its diet (viz. breakdown of cell wall material; detoxification). For instance, there is no sacculation of the stomach as is found in the specialized leaf-eating monkeys of the Old World (Colobinae), and the caecum is not notably enlarged. Using carefully-collected quantitative data on both behaviour and ecology, Milton sets out to define the "foraging

strategy" of the howlers and to explain the lack of obvious digestive tract specialization. In a clear and economically written account, she gradually constructs a picture of the mantled howler as a species which neatly balances considerable selectivity in feeding against tight budgeting of energy spent in travel. It emerges that the BCI howlers preferentially feed on relatively clumped fruit and flower sources when available and complement their diet by selectively feeding almost exclusively on young leaves (which, as a rule, contain more protein and less cell-wall material than mature leaves). Seasonal items are particularly important and the howler modifies its feeding pattern over the annual cycle, feeding most on leaves (including a small proportion of mature leaves) when other items are scarce. Despite its concentration on food items with a spatially and temporally patchy distribution, *Alouatta* is a relatively sluggish primate, resting for about 65% of daylight hours. It seems to save energy by travelling directly to a few "pivotal" trees (usually fruit-sources) each day, picking up the rest of its food with little extra travel. A wide range of different plant species (at least 73 per troop) is exploited and the overall effect is to limit travel distance to the low level of 443m per day. In line with many other primate species, the howler is most active, and has the most diversified diet, during the season of greatest availability of preferred foods. Surprisingly, although the vegetation on BCI is dense tropical rainforest, there is a marked seasonal pattern (consistent from year to year). Still more surprisingly, it is the "dry" season on BCI that is apparently most favourable in terms of availability of preferred foods of the howlers.

A brief review cannot do justice to the theoretical scope and importance of this monograph. In addition to the central theme of foraging strategy, the text broaches a number of important areas — such as the relation between social organization and feeding ecology, the co-evolution of plants and animals and the evolution of plant secondary compounds as defence systems — which are at the spearhead of current primate field studies. This is an outstanding book, with only a few minor shortcomings (for example inadequate discussion of problems of bias in data collection and insufficient integration of the results with information from other quantitative studies of howlers and other primates), and it deserves to be widely read by zoologists, primatologists and anthropologists. As a companion volume to Pierre Charles-Dominique's *Ecology and Behaviour of Nocturnal Prosimians* (1977), Milton's monograph sets a very high standard indeed for any further books on this topic from Columbia University Press. □

R.D. Martin is Reader in Physical Anthropology at University College and Visiting Professor in Zoology at Birkbeck College, University of London.

Is there nothing new in primatology?

Robert A. Hinde

The Social Life of Monkeys and Apes, 2nd Edn. By S. Zuckerman. Pp.511. ISBN 0-7100-0681-8. (Routledge & Kegan Paul: 1981.) £17.50, \$55.

IN 1932, Zuckerman's *Social Life of Monkeys and Apes* was far ahead of its time. It contained scholarly reviews of animal sociology and of the physiology of the oestrous and menstrual cycles. It was remarkable in attempting to integrate physiology with behaviour. Several chapters reviewed the author's own pioneering observations of baboons in South Africa and Hamadryas baboons in the London Zoo. From these he drew conclusions about monkey societies, and about the nature of social and sexual responses, which stimulated much research. Re-reading it now one marvels at the scope of the data synthesized, at the keen critical sense of the author and the number of issues, now the subject of detailed research, which appear there in embryonic form. To mention just one, the importance of "contact comfort" for non-human primates, re-discovered by Harlow in 1959, was discussed by Zuckerman in 1932.

The book stands as a landmark because it helped usher in a new way of studying monkeys. Of course we now know much more about both endocrinology and non-human primates, and new theoretical perspectives have developed. This unrevised re-issue of Zuckerman's book (the addition of a postscript and seven appendices containing reprinted essays or reviews hardly justifies the label of "Second Edition") provides an opportunity to assess the progress that has been made. Three generalizations may serve as examples. First, Zuckerman held that the nucleus of all primate societies is "the family party, consisting of an overlord and his harem" (p.314). Whilst true of hamadryas, recent field studies have shown that some species are monogamous (e.g. gibbons, marmosets), and that superficially polygynous ones show a great diversity of structural types. Whilst in some species there are close relationships between particular males and particular females, in others male-female liaisons are much more temporary (e.g. baboons, macaques).

Second, in 1932 Zuckerman regarded sex as the main force that holds groups

together (e.g. pp.59, 314). But whilst the primary bonds in some polygynous species are between male(s) and females (gorilla), more usually they are between mothers and their daughters. In the latter case, each troop may consist of one or more matriline (a matriarch, her daughters and their daughters) to which males become temporarily attached (macaques, baboons), or are quite peripheral (patas). Zuckerman's emphasis on sexual bonds as paramount seems to stem from a view that single sex groups do not have bonds (pp.58-59), perhaps based on a presupposition that the "family", containing male(s) and females, is necessarily central to any society. The third issue is Zuckerman's emphasis on male dominance as the major principle determining social structure. Dominance is of course important, but we now know that females can be dominant to males for at least some of the time (talapoin, squirrel monkey), and that social structure depends also on other issues, including a correlate of genetic relatedness. Indeed field primatology is now in an exciting stage, with new theoretical insights matched by field data revealing new subtleties in social structure. Primatology, for long lagging behind other branches of biology, is now generating important general principles.

That the book now needs amendment is no reflection on the first edition: we see further now in part because we stand on the shoulders of giants, and for a related reason must also acknowledge that some of our own conclusions may later need extension or correction. But the unfortunate postscript to this re-issue brings no such feeling of humility about the progress of science. It conveys instead the impression of an attempt to discredit not only whole areas of research developed since 1932 but also the contributions of young workers within those fields, and to show that primatology has made little progress in 50 years. In doing so, Zuckerman goes to rather remarkable lengths. He attacks quantification and in particular the field workers' practice of stating the extent of their data base (p.363) and use of statistical techniques (pp.370-371), apparently in part because he mistakenly perceives monkey behaviour to be almost infinitely variable (pp.371, 375-377, 403). He ignores most experimental work, and descends to describing most field workers as "novices with little experience of scientific research or method", incidentally distorting a semi-popular review of a sketchy article on the problems they face (p.366). Some of his criticisms are based on citations which inaccurately represent the original source (e.g. p.371, lines 33-39). He misunderstands functional questions, rejecting research into ultimate causes as ending "either in

empty statement or in tautology" (p.404) and, although earlier regretting the lack of ecological data on primates (p.31), he labels most subsequent research as "tortuous and intellectually unrewarding" (p.375). His emphasis on the variability of monkey behaviour leads him almost to deny inter-species differences (pp.375, 379): he ignores most modern evidence on the diversity of social systems, and tries to discredit the rest. He seems unaware of the advances consequent upon the kin selection hypothesis.

Now of course not all primatology (nor indeed all endocrinology) has been of the highest quality, and apparently promising theoretical alleys can prove unrewarding. Some criticisms, especially those directed at the popular writing of Morris, Ardrey and sometimes Lorenz, are justified. But Zuckerman's postscript gives the impression of being almost wholly destructive, and contrasts with the original book in not attempting a synthesis. This is a pity: the re-issue of this once influential book might have been accompanied by a constructive appraisal so that it, like the original, facilitated future progress. □

Robert A. Hinde is a Royal Society Research Professor and Honorary Director of the Medical Research Council Unit on the Development and Integration of Behaviour, Madingley, Cambridge.

Star of the Galaxy

D. Lynden-Bell

Oort and the Universe: A Sketch of Oort's Research and Person. Edited by Hugo van Woerden, Willem N. Brouw and Henk C. van de Hulst. Pp.210. ISBN hbk 90-277-1180-1; ISBN pbk 90-277-1209-3. (Reidel: 1980.) Hbk \$29, Dfl.55; pbk \$12.95, Dfl.25.

OORT was the heir of a great Dutch tradition in astronomy but under his guidance it has flowered to an extent unforeseen by his predecessors. In this book we have a sketch of Oort's major contributions to galactic and cometary astronomy and catch a glimpse of a man whose foresight and quiet insistence gave Holland the world's most powerful radio telescope at Westerbork.

The book is an eightieth birthday tribute to Oort from 25 leading astronomers who have worked with him or have been close colleagues. Some concentrate on Oort's scientific influence, some on the political difficulties that Oort overcame to get the money for a major instrument and some on the science that it opened to exploration.

However, this is not a *Festschrift* in which leading scientists contribute samples of their work in honour of the recipient; rather, it is a survey of history in the making, concentrating on Oort's period

Journals supplement

On October 1 *Nature* will publish a review supplement devoted to new journals which have appeared between January 1978 and May 1980. Publishers and societies are invited to submit sample copies for consideration for review; for details see *Nature* 291, 599.

and his interests. Here we have some basic material from which a historian might begin his study of the way the character of a great scientist allows him to further science through his influence as well as his own research. Anyone who has studied galactic astronomy, radio astronomy or cometary astronomy will find here some fascinating pieces of history. B. Lindblad's picture of galactic rotation was that the stars with greater velocity dispersion had slower rotation about the galaxy because of their greater pressure support. Each such subsystem was originally considered to rotate uniformly. It was Oort who saw that stars of very low velocity dispersion must rotate differentially and who then went on to show that this motion could already be seen in the observed proper motions.

Later, it was Oort who saw the importance of finding spectral lines in the radio region and sent van der Hulst on the search that led to his great discovery of the 21 cm line. Until reading this book I had always regarded it as a sad irony of fate that the Dutch, with that prediction but with no wartime experience of radar, just missed being the first to observe the line. Much credit goes to Muller for bringing the techniques up to date so quickly, but the gap was too great and, as he points out here, his task was made quicker by learning of the techniques used in the US which led Ewen to the first detection. However, Oort knew what should be done with the line to probe the astronomy of the Galaxy and so it was the Dutch who won the next round. It is sad that G. W. Rougoor who collaborated with Oort in some striking papers died when so young, but I am sure he would too have been a contributor.

However, if I may choose a favourite among Oort's discoveries, it is his determination of the total density of visible and invisible matter in our region of the Galaxy through his work on the gravity field. As Blaauw describes, this work opened the door to the astronomy of the invisible which has grown to be a major subject over the past ten years and in which the Westerbork telescope has played a major part. Some of the highlights of the discoveries are described in the well-illustrated contribution by Allen and Ekers, while van der Laan describes Oort's part in the use of Westerbork observations in his new role as Professor Emeritus.

I give the last word to van Bueren who contributes a well-written and perceptive article on Oort's role in Dutch astronomy: The kind and courteous Oort holds fast to his opinions with unbelievable firmness. He therefore finally always wins his argument, with colleagues as well as with research-financing organisations and government officials.

Perhaps this has been Oort's secret of success, which has helped him to put Holland in the forefront of modern astronomy. □

D. Lynden-Bell is Professor of Astrophysics at the Institute of Astronomy, Cambridge.

Rheology from Russia

J.J. Hermans

Rheology of Polymers: Viscoelasticity and Flow of Polymers. By G.V. Vinogradov and A. Ya. Malkin. Pp.467. ISBN 3-540-09778-3. (Springer-Verlag: 1981.) \$57.90, DM 98.

THIS translation of a revised Russian text published in 1977 contains an admirable treatment of the rheology of polymers. It begins with a comprehensive introduction to the mathematics of deformation and flow, which covers both linear and nonlinear constitutive equations.

More than 30 pages on linear viscoelasticity are followed by rather brief, but pertinent, coverage of nonlinear theories. There are excellent sections on the newtonian and the non-newtonian viscosity of polymers in relation to molecular weight, polydispersity, branching, concentration and entanglements. A remarkably extensive discussion of about 40 pages covers normal stresses. There are further chapters on birefringence, rubber-like behaviour in flow (i.e. large recoverable deformations), the effect of fillers and the rheological properties of polymer blends.

The authors have wisely decided to emphasize concentrated solutions and melts, these being the systems of greatest practical interest. The well-known molecular models and theories that have been developed for very dilute polymer solutions are mentioned, but play no great role in the discussion of either the theory or the experimental data. In other words, the treatment is mostly phenomenological. Many empirical equations are discussed but, when at all possible, these are related to the phenomenological theory.

Particularly praiseworthy is the balance between theory and experiment. This is perhaps best illustrated by the fact that Chapters 2-7 contain about 180 figures showing experimental results.

Inevitably, in a book which covers such a vast field, one will notice omissions. For example, whereas empirical relations between non-newtonian viscosity and shear stress are mentioned, there is nothing on the requirement of zero slope at the limit of zero shear stress. To mention another example, in the chapter on uniaxial extension the investigator will find that many literature references are missing; conversely, it can be said that the value of the book is enhanced by some references with which the English, and even more so the North American reader is often unfamiliar.

The book will be a good introduction for anyone wishing to learn about this field. At the same time it will be a valuable reference work to scientists already experienced in the subject. □

J.J. Hermans is Emeritus Professor of Physical Chemistry at the University of North Carolina.

SCIENTIFIC BOOKSHOP

H.K. LEWIS can supply works in all branches of Pure and Applied Science. Catalogues on request. Please state interests.

SCIENTIFIC LENDING LIBRARY

Annual Subscription from £7.50.
(Available in U.K. only)

Reduced rates for multiple subscriptions.

Prospectus post free on request.

Quarterly List of New Books and new editions added to the Library sent post free to subscribers regularly.

**H.K. LEWIS
& Co. Ltd.**

136 GOWER STREET,
LONDON, WC1E 6BS

Telephone: 01-387 4282
Telegrams: "Publicavit,
London, WC1E 6BS."

Circle No. 14 on Reader Enquiry Card.

**Bring Your
Library Up-To-Date
With This ANS Edition...**

**J.M. Cleveland
MONOGRAPH**

The Chemistry Of Plutonium

Approx. 680 pages 6" x 9"
Hardbound \$49.00

**American Nuclear Society
555 North Kensington Avenue
La Grange Park, IL 60525 USA**

Please take my order for _____ book(s).

☐ Payment in full enclosed.

☐ Bill me and I accept postage and handling charges.

Name _____

Organization _____

Street _____

City _____

State _____ Zip _____

Circle No. 02 on Reader Enquiry Card.

Social scientists and sociobiologists get their lines crossed

Jon Seger

Sociobiology Examined. Edited by Ashley Montagu*. Pp.272. ISBN hbk 0-19-502711-6; ISBN pbk 0-19-502712-4. (Oxford University Press: 1980.) Hbk £12.50, \$19.95; pbk £2.95, \$5.95.

Sociobiology Examined is a collection of 16 articles, half of which have been or will be published elsewhere. The authors are listed below. According to the dust jacket, "the articles . . . examine the claims [E.O.] Wilson makes for [human] sociobiology [in *Sociobiology* and *On Human Nature*] and, by a critical examination of the evidence (or lack of it) on which these claims are based, assess their validity". In only a few of the articles is there a sustained attempt to do this. Some are mostly polemical, in places irresponsibly so; some are mostly about subjects other than sociobiology. For example, Layzer's article turns out to be a detailed critique of Piaget's theory of cognitive development. Sheehan's is a review of recent books on political philosophy by Bernard Henri-Levy and Alain de Benoist, neither of whom appears to have the slightest idea what sociobiology is or says, except that someone told them it supports their opinions. Most of the articles stick more closely to the stated subject than do these, but taken as a whole they give the impression that "sociobiology" is more of a Rorschach test than it is a theory or even a well-defined point of view. Midgley comes to a similar conclusion in her essay on the first few years of the "debate" on sociobiology. Although the articles present a bewildering range of views, two major themes come up fairly often. One concerns the supposed social, political or ethical implications of sociobiological theories, and the other the issue of genetic determinism.

In agreement with Plato, Hobbes, Rousseau and Wilson, most commentators on sociobiology seem to believe that a theoretically and empirically adequate description of human nature necessarily implies a morally compelling prescription for the organization of human society. The lone dissenter in *Sociobiology Examined* is Mackintosh, whose article was first published as a review of *On Human Nature* (*Science* 204, 735; 1979). "The gap between 'is' and 'ought' remains as wide today as it was when Hume pointed it out 200 years ago". Mackintosh quotes Wilson's claim that "a correct application

of evolutionary theory . . . favors diversity in the gene pool as a cardinal value", and notes that this is only one of several places where Wilson states or implies that a value can be derived from a fact. But as Mackintosh also notes, Wilson's "radical" critics appear to be even more committed to the naturalistic fallacy than he is. It is truly incredible that in six years of heated "debate" there has been hardly any serious discussion of this problem.

The question of genetic determinism is much more complex. As Midgley points out, "determinism" is simply a belief in the efficacy of causes. Thus unless you are a dedicated supernaturalist, or believe in at most one cause, it is silly to accuse someone of being a "determinist". As a matter of logic this is true enough, but the social scientists appear to have something more concrete in mind when they speak of genetic or biological determinism. Unfortunately, few of the authors in *Sociobiology Examined* ever state explicitly what they mean by phrases like "genetic basis", but by the end of the book it has become clear that many of them believe sociobiology attempts to map differences of behaviour, whether between individuals or societies, onto corresponding differences of genotype. In other words, any variability that has a "biological" explanation is necessarily heritable. This belief can be seen clearly in the article by Simon.

Behavior that can plausibly be designated as innate or instinctive must be stable throughout a range of environments . . . Human social behavior . . . varies considerably from one social group to another. Social science is typically interested in explaining these variations . . . If the determinants of the specific differences among a widely disparate range of behavior are not biological but sociocultural, it will not explain any of them to subsume them all under a common biological rubric.

Sociobiology does, in part, attempt to explain behavioural variation. But does it really argue for the *heritability* of such variation? Those who think that sociobiological theory derives from traditional behaviour genetics say "yes"; this group includes most of the authors in the present collection, but very few sociobiologists. Those who identify sociobiology with the newer work on sex ratios, sexual dimorphism, kin-directed altruism, life-history strategies and the like, say "no"; this group includes most sociobiologists, but in the present collection only Barkow, who notes that "Evolutionary biology does not require . . . a direct link . . . between a particular gene and a particular behavior, even if careless reading (and writing) of sociobiology may occasionally permit that impression". There is no contradiction in this as long as different levels of causation

are properly distinguished. Barkow's discussion of the levels of causation needed to explain human behaviour is well worth reading. He observes, as have many others, that selection may often favour behavioural flexibility of a kind "mediated by learning preferences". There would be little point to flexibility unless it enabled an individual to recognize and seize unique opportunities. Consistent with this general expectation, animals in learning experiments tend to master some associations much more easily than others. If sociobiology is really about the evolution of learning and other environmentally mediated developmental processes, then at the level of differences between individuals it is, ironically, a theory of *environmental* determination. Individual differences map onto differences of environment, not of genotype, although they do so according to a logic that is "genetic" in the sense that it reflects innate, evolved characteristics of the species. Thus only at the level of differences between species is sociobiology inherently a theory of genetic determination. Given these contrasting views of "sociobiology", and given the distinctions on which they turn, any simple reference to a genetic substrate, basis, cause, determination or the like, unless carefully specified, is almost certain to be literally meaningless. Because these key terms are seldom specified in the pages of *Sociobiology Examined* the reader often cannot tell just where a given piece of criticism is supposed to strike.

Without a doubt the sociobiologists are partly responsible for this widespread confusion regarding levels of causation and explanation. Mackintosh rightly criticizes Wilson for shifting between different senses of "the notion of genetic determination", and more generally for not giving us "a serious and sustained attempt to analyze in just what sense our behavior or social organization is genetically determined . . .". But if Wilson's writing is careless or confused about something so fundamental, why should that of his sharpest critics be even more so? Surely critics have a responsibility to expose weaknesses, rather than to hide them.

Several observers of the sociobiology controversy have remarked that the participants seem to be talking past each other. Unfortunately, *Sociobiology Examined* does more to document this situation than to change it. The book will be of interest to historians and sociologists of science. But for the most part it will be of little help to those who are still wondering how the study of animal behaviour may ultimately improve the study of human behaviour, and vice versa. □

Jon Seger is at the Biological Laboratories, Harvard University.

* Articles by Jerome H. Barkow, S.A. Barnett, Derek Freeman, Stephen Jay Gould, Marvin Harris, James C. King, David Layzer, N.J. Mackintosh, Mary Midgley, Ashley Montagu, Karl Peter and Nicholas Petryszak, Stephen Rose, Thomas Sheehan, Michael A. Simon and S.L. Washburn.